MARYLAND

TECHNICAL
RESEARCH
REPORT

SRC TR 87-213

Some Solutions to Problems in Depth Vision from Differential Geometry and Deconvolution Methods

by

E.V. Patrick

SYSTEMS RESEARCH CENTER

UNIVERSITY OF MARYLAND

COLLEGE PARK, MARYLAND 20742

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **1987**	2. REPORT TYPE	3. DATES COVERED **00-00-1987 to 00-00-1987**

4. TITLE AND SUBTITLE **Some Solutions to Problems in Depth Vision from Differential Geometry and Deconvolution Methods**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland at College Park,Systems Research Center,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Two distinct parts are presented. The first part is an application of manifold theory and geometry to mathematical modeling problems in depth vision. The second part is an application of deconvolution methods to the problem of constructing converging sequences of approximating functions from sampled values of convolutions. The connection between the two parts is this: the first part depends on differential methods; the second part provides converging algorithms for such methods. The first part begins with a model for objects. A measure is introduced and smoothness is assumed almost everywhere. The radiometric notion of sterance is modeled using differential forms on the sphere bundle (of three dimensional Euclidean space). It is shown that sterance, even if known on a neighborhood in the sphere bundle does not uniquely determine the objects. However, sterance on a neighborhood can be used to construct a submersion, hence to determine codimension one submanifolds. A similar construct ion is carried out for sterance given on the sphere bundle over a curve (motion stereo with the path known). The properties of this construction are used to motivate the constructions for modeling the problem of depth vision based on time varying sterance on the sphere bundle over a point. We introduce bundles of bases, integral manifolds, and vector fields and one forms with vanishing Lie derivative with respect to the position vector field. We study the flows of vector fields that are isometries of the integral manifolds. The second part begins with an analysis of deconvolution methods for convolution operators that are characteristic functions of n-dimensional cubes. Such operators (for squares) are approximations of the impulse response of photo-detectors in . vision systems. A complete description for the implementation of the method for sampled data is given. The primary accomplishment is that the error analysis is presented with explicit error bounds throughout. The final chapter is an analysis of the properties of the deconvolution methods in the presence of additive noise. For the class of methods and noise studied it is shown that there is no penalty for the use of deconvolution methods with photo-detectors.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**246**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

ABSTRACT

Title of Dissertation: Some Solutions to Problems in Depth Vision from Differential Geometry and Deconvolution Methods

Emil Vincent Patrick, Doctor of Philosophy, 1987

Dissertation directed by: James C. Alexander, Professor
and
Carlos A. Berenstein, Professor
Department of Mathematics

Two distinct parts are presented. The first part is an application of manifold theory and geometry to mathematical modeling problems in depth vision. The second part is an application of deconvolution methods to the problem of constructing converging sequences of approximating functions from sampled values of convolutions. The connection between the two parts is this: the first part depends on differential methods; the second part provides converging algorithms for such methods.

The first part begins with a model for objects. A measure is introduced and smoothness is assumed almost everywhere. The radiometric notion of sterance is modeled using differential forms on the sphere bundle (of three dimensional Euclidean space). It is shown that sterance, even if known on a neighborhood in the sphere bundle, does not uniquely determine the objects. However, sterance on a neighborhood can be used to construct a submersion, hence to determine codimension one submanifolds. A similar construction is carried out

for sterance given on the sphere bundle over a curve (motion stereo with the path known).

The properties of this construction are used to motivate the constructions for modeling the problem of depth vision based on time varying sterance on the sphere bundle over a point. We introduce bundles of bases, integral manifolds, and vector fields and one forms with vanishing Lie derivative with respect to the position vector field. We study the flows of vector fields that are isometries of the integral manifolds.

The second part begins with an analysis of deconvolution methods for convolution operators that are characteristic functions of n-dimensional cubes. Such operators (for squares) are approximations of the impulse response of photo-detectors in vision systems. A complete description for the implementation of the method for sampled data is given. The primary accomplishment is that the error analysis is presented with explicit error bounds throughout.

The final chapter is an analysis of the properties of the deconvolution methods in the presence of additive noise. For the class of methods and noise studied it is shown that there is no penalty for the use of deconvolution methods with photo-detectors.

SOME SOLUTIONS TO PROBLEMS IN DEPTH VISION FROM DIFFERENTIAL GEOMETRY AND DECONVOLUTION METHODS

by

Emil Vincent Patrick

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfullment
of the requirements for the degree of
Doctor of Philosophy
1987

Advisory Committee

Professor James C. Alexander
Professor Carlos A. Berenstein
Professor Karsten Grove
Associate Professor P. S. Krishnaprasad
Assistant Professor Eyad Abed

ACKNOWLEDGMENTS

This work was sponsored in part by the Systems Research Center of the University of Maryland under grants NSF-CDR-8500108 and NSF-DMS-8703072, and in part by the Center for Night Vision & Electo-Optics, Fort Belvoir, VA. The support of these two organizations is gratefully acknowledged. Special thanks go to my two coadvisors, James Alexander and Carlos Berenstein, first for agreeing to be coadvisors and second for successfully doing it.

CONTENTS

INTRODUCTION

This work is the product of the confluence of several technical disciplines at a particular problem topic. Each of the disciplines, as well as the problem topic, is well established. What is unusual is the drawing together of the particular collection of disciplines about the problem. As the title suggests, the blending is unfinished: the work continues along two main branches but has not yet merged.

The objective of this introduction is to provide some background and rationale for our particular collection of techniques and problem focus.

First the problem. The problem area is vision. But this needs both some qualification and expansion. (For an introduction to contemporary issues and progress in understanding and synthesizing vision phenomena see Marr (1982).) We are interested in the subtopic of vision that addresses the remote sensing of objects as subsets of three dimensional Euclidean space. This should be contrasted with the subtopics of vision that are solely pattern recognition, for example, reading and photo-interpretation. The subject of remote sensing of objects really extends beyond vision. It includes such things as radar, thermal sensing, laser radar, sonar, and tomography. In fact, the problem area is possibly more accurately described by the term space perception. However, at least in vision, the most widely

recognized term is depth vision. As this work addresses only vision questions, we shall use this descriptor. Collett and Harkness (1982) provide an excellent survey of the experimental knowledge of the variety of depth vision phenomena exhibited by animals.

The aspect of depth vision that is of interest in this work is mathematical modeling and the questions of existence and uniqueness of solutions associated with the models. This raises the question of what is a solution. It has long been my bias that any plausible mathematical model for depth vision must have manifolds as solutions. Thus, manifold theory and all of the related mathematical techniques are a primary element in our collection. Throughout this work the standard problems in which the solutions are manifolds were kept in mind: submersions, transversal intersections, and integration of involutive distributions.

A second element in our collection of techniques is measure theory. We use both Lebesgue and Hausdorff measure to make precise such issues as edges and corners and to provide a weaker notion of manifold.

A related issue is functions. Smooth functions are generally regarded as too narrow a space for modeling. Since smooth functions and differentiable structure are essentially equivalent, we must be careful in the choice of function space. Consequently, in everything we do here we understand that both the manifolds and the functions are smooth approximations to objects in a more general space. In other

words, we are free to use mollifiers, i.e., convolution, as needed.

This point of view is not only mathematically useful but is also fruitful in modeling, for many of the physical processes in vision are indeed (approximately) convolutions. But this raises the issue of approximation and convergence. Our method to resolve this dilemma is the use contemporary methods of deconvolution. This topic is a further major element in our collection of techniques. We use deconvolution to provide both a theoretical and practical tool to obtain well defined converging approximating sequences for the manifolds and functions in our model. It is this convergence in a suitable function space that justifies our use of differential methods.

The manifold methods and the deconvolution methods are not yet fully merged. As of now the two are developing along separate, parallel lines. Their individual progress is the subject of this document.

PART 1

SOME SOLUTIONS TO PROBLEMS IN DEPTH VISION FROM DIFFERENTIAL GEOMETRY

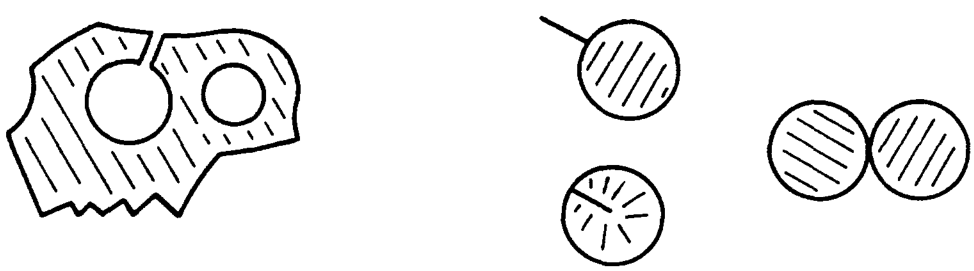
1 SOLUTIONS IN THE SPHERE BUNDLE

1.1 Objects, Hausdorff measure, and smooth almost everywhere

In this first section a somewhat technical issue is addressed. The first goal is to provide a sufficiently careful mathematical description of precisely what it is one is attempting to solve for in a depth vision or visual perception problem. That is, in this section an answer is provided to the questions of what constitutes a suitable definition of "objects" and what are the consequent properties of this definition. The motivation for all of this is contained in Theorem 8. Theorem 8 may be paraphrased by saying that differential methods may be applied on an open, dense subset and that the complement of the subset has measure zero. Paraphrased in terminology borrowed from computer vision, Theorem 8 says that all of the so called "edges" and discontinuities are contained in a closed set of measure zero.

1. DEFINITION. An *object* A is a compact, connected subset of $\mathbb{R}^3$ that is a topological 3-dimensional submanifold with boundary, wherein the boundary ∂A contains an open set U which is a smooth 2-dimensional submanifold, and the 2-dimensional Hausdorff measure of $\partial A - U$ is zero.

This definition is made, as usual, with the objective of including at least those things of interest while excluding as much as possible of all else. This objective is illustrated in Figure 1. (In Figure 1, as well as in others that follow, two dimensional figures are used to illustrate higher dimensions.) In this definition the subset U is a submanifold in the sense that U is a smooth manifold with the subset topology and the inclusion map is an immersion (Bishop and Goldberg 1968, §1.4). See the proof of Theorem 8 for the definition of Hausdorff measure (Federer 1969; Evans and Gariepy lecture notes).



object

not objects

Figure 1

It is certainly desirable to address the possibility of more than one object. A way to do this is provided by the next definition.

2. DEFINITION. A *set of objects* is a finite collection $\{A_i\}_{i=1}^{N}$, where N is a positive integer and each A_i is an object, such that

i) the interiors of the objects are pairwise disjoint,

$$\text{Int } A_i \cap \text{Int } A_j = \emptyset \text{ for } i \neq j ,$$

ii) the union of the intersections, $\bigcup_{i\neq j}\left[\,\partial A_i \cap \partial A_j\,\right]$, and the boundary of the union, $\partial\left(\bigcup_i A_i\right)$, have an intersection which has a 2-dimensional Hausdorff measure of zero.

We shall denote the union of all objects by $A = \bigcup_i A_i$.

The motivation behind this definition is illustrated in Figure 2.

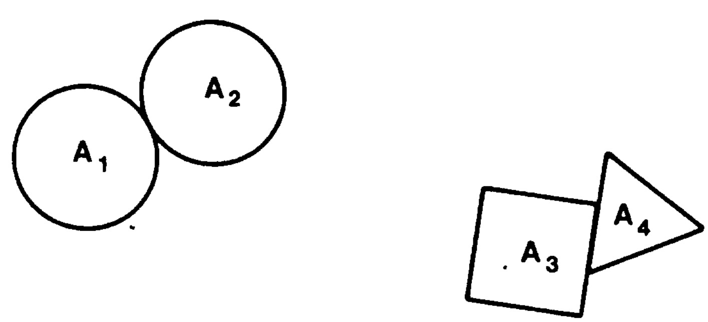


Fig. 2. A set of objects

It is clear that all of the following are compact: A, ∂A, ∂A_i, $i=1,\ldots,N$, and $\bigcup_i \partial A_i$. Hence, $\partial A \cap \bigcup_{i\neq j}\left[\partial A_i \cap \partial A_j\right]$ is closed, and it has a 2-dimensional Hausdorff measure of zero by definition. For each ∂A_i, $i=1,\ldots,N$, let U_i be the (relatively) open set in ∂A_i which is a smooth submanifold. Then $\partial A \cap \bigcup_i\left[\partial A_i - U_i\right]$ is closed and has 2-dimensional Hausdorff measure zero. It follows, by deleting these two closed (in $\mathbb{R}^3$) sets from ∂A, that ∂A contains an open set U which is a smooth 2-dimensional submanifold of $\mathbb{R}^3$ and $\partial A - U$ has 2-dimensional Hausdorff measure of zero.

The purpose of Definition 2 is not only to permit the consideration of more than one object but also to permit these objects to be in contact. In permiting contact it is necessary to drop the requirement that the boundary be a topological manifold: see, for example, the two spheres in contact in Figure 2. However, the second half of Definition 1 can be retained: up to a closed set of measure zero the boundary is a smooth submanifold. A further remark on Definition 2 is that for many of the situations that follow, the finiteness of the collection could be relaxed to, say, countability with finiteness on any compact region. Essentially what is needed is a suitable analog for the notion of σ-finiteness with respect to a measure. This extra generality in not pursued here.

This discussion began with the definition of objects only because it is necessary to define what ultimately is to be excluded. Nearly all of the analysis is performed on the complement of the objects.

3. DEFINITION. *Empty space* M is the complement of the union of all objects,

$$M = \mathbb{R}^3 - A = \mathbb{R}^3 - \bigcup_i A_i \ .$$

Consequently, M is an open submanifold of $\mathbb{R}^3$.

The usual notation is used for the following objects. $T\mathbb{R}^3$ and TM are the tangent bundles of $\mathbb{R}^3$ and M respectively. The bracket notation $\langle , \rangle$ is used for the standard metric on $T\mathbb{R}^3$ and TM. The sphere bundles over $\mathbb{R}^3$ and over M are denoted by $S\mathbb{R}^3$ and SM respectively, these being the bundles consisting of unit tangent vectors. The projection map of a bundle to its base is usually denoted by π, for example, the projection $\pi: SM \longrightarrow M$.

All of the bundles above are parallelizable. For example, SM is bundle equivalent to $M \times S^2$, that is, there exists a diffeomorphism which maps fibers onto fibers and is the identity on the base. The restriction of a bundle to a subset of the base is denoted by a subscript. For example, the restriction of $S\mathbb{R}^3$ to $\partial A \subset \mathbb{R}^3$ is denoted by $S_{\partial A}\mathbb{R}^3$. Of course, since M is open $S_M\mathbb{R}^3 = SM$, and the latter notation will be used here.

A method is needed for extending a function (or section) from a small neighborhood to some larger open set. In this work we tackle the simplest possible problem and choose the simplest possible method of extension. Roughtly speaking, we use the model that light travels in straight lines in M. The following definition and constructions

exploit this choice. They are illustrated in Figure 3.

The convenient way to formalize the straight line extension is to use geodesic phase flow (Sternberg 1983, 199; Arnold 1980, App.1.J). The standard construction is the following (O'Neill 1983, 67-70). First, the so called natural covariant derivative is defined for vector fields on $\mathbb{R}^3$ by the condition that the standard basis for $T\mathbb{R}^3$ determined by the natural coordinates is parallel. A geodesic is a curve whose velocity vector field along the curve is parallel. Geodesics exist at least locally and are uniquely determined by the initial velocity vector. For $\mathbb{R}^3$ a geodesic, a curve from an interval in $\mathbb{R}$ to $\mathbb{R}^3$, can be extended to the so called maximal geodesic so that it is defined for all of $\mathbb{R}$. Also, the norm of the velocity vector of a geodesic is constant along the geodesic. These standard results are exploited in the following notation.

Let $u \in S\mathbb{R}^3$. Let $\gamma_u : \mathbb{R} \longrightarrow \mathbb{R}^3$ be the maximal geodesic such that the initial velocity vector $\dot{\gamma}_u(0)$ satisfies $\dot{\gamma}_u(0) = u \in S\mathbb{R}^3$.

This formalism is convenient for a number of reasons. First and foremost, it provides a coordinate free represenatation. Second, it lifts in the appropriate way to paths in $S\mathbb{R}^3$. And third, by virtue of this notation everything is in place to consider the problems discussed here for more general, nonflat manifolds.

It is now necessary to use some less standard notation. Here the relation between distance and the parameterization of unit speed geodesics is exploited to define a function on the sphere bundle over

empty space M.

4. DEFINITION. Let SM be the sphere bundle of $\mathbb{R}^3$ restricted to empty space M and let $\mathbb{R}_+ = \{ t \in \mathbb{R} \mid t > 0 \}$. Define $\tau : SM \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ by $\tau(u) = \sup\left\{ t > 0 \mid \forall\, s \in [0,t) : \gamma_u(-s) \in M \right\}$.

We shall refer to the function τ as the *extent* function. Note that in the definition of τ the geodesic curve is followed backwards in its parameter. This is illustrated in Figure 3, where the base point of $u \in SM$ is m, that is, $\pi(u) = m$, and where only $S_m\mathbb{R}^3$ is illustrated.

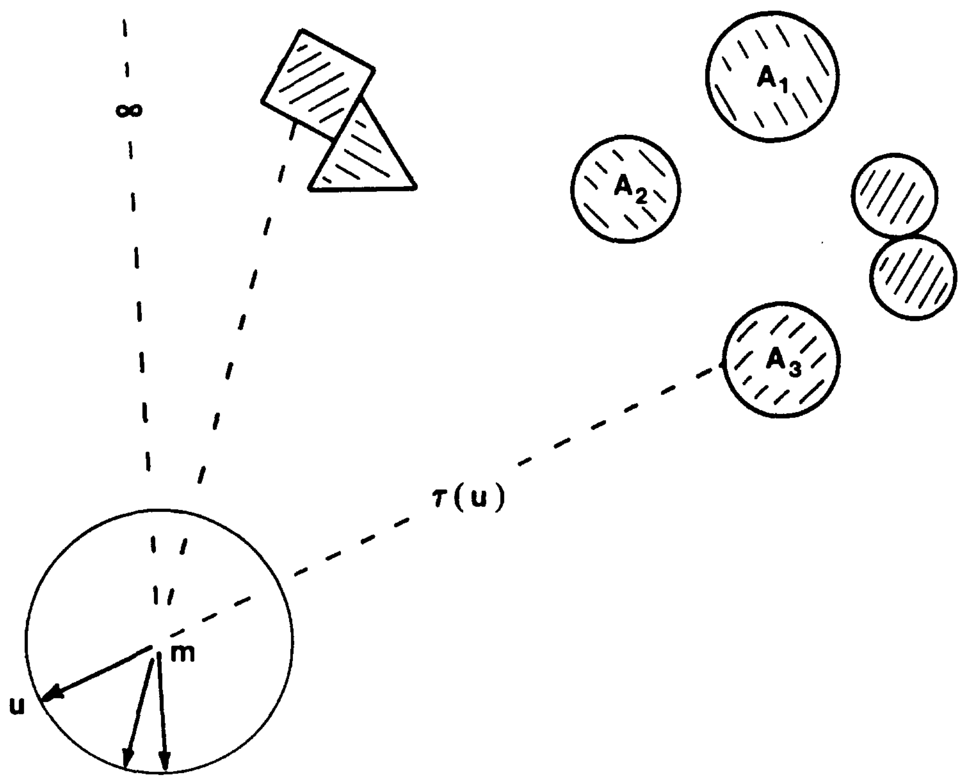


Figure 3

For purposes of notation the following maps are defined:

$$\gamma : S\mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}^3 , \qquad \gamma(u,t) = \gamma_u(t) ;$$
$$\psi : S\mathbb{R}^3 \times \mathbb{R} \longrightarrow S\mathbb{R}^3 , \qquad \psi(u,t) = \dot{\gamma}_u(t) ;$$
$$\psi_t : S\mathbb{R}^3 \longrightarrow S\mathbb{R} , \qquad \psi_t(u) = \psi(u,t) .$$

The map ψ is the so called geodesic phase flow. The map ψ is smooth since ψ is the solution to a smooth ordinary differential equation. Recall that for t, $t' \in \mathbb{R}$, $\psi\big(\psi(u,t) , t'\big) = \psi(u,t+t')$.

The visual perception problem in which we are interested may be stated as the problem of determining the extent of empty space about an observation point. In terms of the structure given above this problem is the determination of the map Φ defined as follows.

5. DEFINITION. Let SM and $S_{\partial A}\mathbb{R}^3$ be the sphere bundle $S\mathbb{R}^3$ restricted to free space M and to the boundary of the union of the objects ∂A, respectively. Let τ be the extent function and let $\mathbb{R}_+ = \{t > 0\}$. Define

$$\Phi : SM - \tau^{-1}(\infty) \longrightarrow S_{\partial A}\mathbb{R}^3 \times \mathbb{R}_+ ,$$
$$\Phi(u) = \Big(\psi\big(u, -\tau(u)\big) , \tau(u) \Big) .$$

The next task is to clarify some properties of the extent function.

6. LEMMA. The extent function τ is lower semicontinuous.

Proof. By definition τ is lower semicontinuous at $u_o \in SM$ if

$$\liminf_{\substack{u \to u_o \\ u \in SM}} \tau(u) \geq \tau(u_o) .$$

Choose t_o such that $0 < t_o < \tau(u_o) \leq \infty$. Hence, for all $s \in [0,t_o]$, $\gamma_{u_o}(-s) \in M$. Let $\gamma: S\mathbb{R}^3\times\mathbb{R} \longrightarrow \mathbb{R}^3$ be defined by $\gamma(u,t) = \gamma_u(t)$. Let A be the union of all objects. By continuity $\gamma^{-1}(A)$ is closed in $S\mathbb{R}^3\times\mathbb{R}$, while $\{u_o\}\times[0,t_o]$ is compact and disjoint from $\gamma^{-1}(A)$. For each point $(u_o,s) \in \{u_o\}\times[0,t_o]$ there exists an open set $U_s \subset S\mathbb{R}^3$ and an $\varepsilon_s > 0$ such that $(u_o,s) \in U_s\times(s-\varepsilon_s,s+\varepsilon_s) \subset \gamma^{-1}(M)$. Passing to a finite subcover $\left\{U_{s_i}\times(s_i-\varepsilon_{s_i},s_i+\varepsilon_{s_i})\right\}_i$ of $\{u_o\}\times[0,t_o]$, we have, with $U_o = \bigcap_i U_{s_i}$, that there exists $\eta > 0$ such that

$$\{u_o\}\times[0,t_o] \subset U_o\times(-\eta,t_o+\eta) \subset \bigcup_i U_{s_i}\times(s_i-\varepsilon_{s_i},s_i+\varepsilon_{s_i}) \subset \gamma^{-1}(M).$$

Consequently, for every $u \in U_o$ we have $\tau(u) \geq t_o + \eta$, hence $\inf_{u\in U_o} \tau(u) \geq t_o + \eta$ and $\liminf_{\substack{u \to u_o \\ u \in SM}} \tau(u) \geq t_o + \eta > t_o$. Since t_o was an arbitrary positive number less than $\tau(u_o)$, the result follows. ∎

7. COROLLARY. $\tau^{-1}(\infty)$ is open in SM.

Proof. Let τ_n be the restriction of τ to $SM \cap SB_n$, where B_n is an open ball of radius n, n=1,2,..., and $SB_n = S_{B_n}\mathbb{R}^3$. Since $A = \bigcup_i A_i$ is compact, the image of τ_n is contained in a set $(0,K_n)\cup\{\infty\}$, for some positive, finite K_n, for the finite values of τ_n are bounded by the

diameter of $B_n \cup A$. Consequently, $\tau_n^{-1}(\infty) = \tau_n^{-1}(\{t>K_n\}) = \tau^{-1}(\{t>K_n\}) \cap SB_n$ which is relatively open in SM since τ is lower semicontinuous and since SB_n is open (Wheeden and Zygmund 1977, Theorem 4.14). Hence $\tau^{-1}(\infty) = \bigcup_n \tau_n^{-1}(\infty)$ is open in SM. ∎

The following theorem is the primary goal of this section. It tells us something about how "edges" of objects and other troublesome sets on the boundary of objects appear to an "observer" in SM.

8. THEOREM. Let the set of objects be non-empty. Then the following hold.

i. The interior of $SM - \tau^{-1}(\infty)$ is non-empty.

ii. The interior of $SM - \tau^{-1}(\infty)$ contains an open set $\mathcal{G}$ on which Φ is a diffeomorphism of $\mathcal{G}$ with its image in $S_{\partial A}\mathbb{R}^3\times\mathbb{R}_+$, and $\mathcal{Z} = SM - \tau^{-1}(\infty) - \mathcal{G}$ has measure zero in the smooth manifold SM.

iii. Let $S(\partial A - Z)$ denote the sphere bundle of $\partial A - Z$. With $S(\partial A - Z)$ identified with its inclusion in $S_{\partial A}\mathbb{R}^3$,

$$\mathcal{Z} = \Phi^{-1}\Big(\big[S(\partial A - Z)\times\mathbb{R}_+\big] \cup \big[S_Z\mathbb{R}^3\times\mathbb{R}_+\big]\Big) .$$

Proof. Recall the map $\psi\colon S\mathbb{R}^3\times\mathbb{R} \longrightarrow S\mathbb{R}^3$, $\psi(u,t) = \dot{\gamma}_u(t)$. We make several observations.

Obs. 1 ψ is a left inverse for Φ, for $\psi\big(\Phi(u)\big) = \psi\big(\psi\big(u,-\tau(u)\big) , \tau(u)\big) = \psi(u,0) = u$.

Obs. 2 Φ is one-to-one, for it has a left inverse.

Obs.3 For any set $S \subset S\mathbb{R}^3\times\mathbb{R}$, $\Phi^{-1}(S) = (\psi\circ\Phi)(\Phi^{-1}(S)) \subset \psi(S)$.

CLAIM 1. For $Z \subset \partial A \subset \mathbb{R}^3$ and the 2-dimensional Hausdorff measure of Z zero, then $\Phi^{-1}(S_Z\mathbb{R}^3\times\mathbb{R}_+)$ has measure zero as a subset of the smooth manifold *SM*.

Proof of Claim 1. By Obs.3 is suffices to show $\psi(S_Z\mathbb{R}^3\times\mathbb{R}_+)$ has measure zero .

The Hausdorff measure is defined for subsets of $\mathbb{R}^n$. To relate $S\mathbb{R}^3 \times \mathbb{R}_+$ to $\mathbb{R}^6$ let $T\mathbb{R}^3-\{0\} = \left\{v \in T\mathbb{R}^3 \mid \langle v,v\rangle \neq 0\right\}$ be identified with $\mathbb{R}^3 \times (\mathbb{R}^3-\{0\})$ by the usual identification using the natural coordinates for $\mathbb{R}^3$. Define $\rho: T\mathbb{R}^3-\{0\} \longrightarrow S\mathbb{R}^3 \times \mathbb{R}_+$ by $\rho(v) = \left(\frac{v}{\|v\|}, \|v\|\right)$, where $\|v\| = \langle v,v\rangle^{1/2}$. Then $\psi\circ\rho$ is smooth, hence locally Lipshitz, on $T\mathbb{R}^3-\{0\}$.

The following result is needed. For any set $A \subset \mathbb{R}^n$, let $H^s(A)$ denote the s-dimensional Hausdorff measure of A. If $H^s(A) = 0$, then, for $A \times \mathbb{R} \subset \mathbb{R}^n \times \mathbb{R}$, $H^{s+1}(A \times \mathbb{R}) = 0$, $0\le s<\infty$. To see this it suffices to consider $s>0$, for H^0 is counting measure. By definition, for $0<\delta\le\infty$, $H^s(A) = \lim_{\delta\to 0} H^s_\delta(A)$, where

$$H^s_\delta(A) = \inf_{\{C_j\}}\left\{ \sum_{j=1}^{\infty} \alpha(s)\left[\frac{\operatorname{diam}C_j}{2}\right]^s \,\middle|\, A \subset \bigcup_{j=1}^{\infty} C_j \,,\ \operatorname{diam}C_j \le \delta \right\}$$

and where $\alpha(s)$ is the Lebesgue measure of a ball with unit radius in $\mathbb{R}^s$. Since $H^s_\delta(A) \ge 0$ and since $H^s_\delta(A)$ increases as δ decreases, $H^s_\delta(A) = 0$ if $H^s(A) = 0$. Fix $\delta < 1/3$ and choose $\varepsilon > 0$. Then there

exists sets $\{C_j\}_{j=1}^{\infty}$ such that $A \subset \bigcup_{j=1}^{\infty} C_j$, $\operatorname{diam} C_j < \delta$, and such that $\sum_{j=1}^{\infty} \alpha(s) \left[\frac{\operatorname{diam} C_j}{2}\right]^s < \varepsilon$. Let I_j^k be the intervals $[k\cdot\operatorname{diam} C_j\ ,\ (k+1)\cdot\operatorname{diam} C_j]$, $k=0,1,\dots,k(j)$, where $k(j) \in \mathbb{N}$ such that $1 \le k(j)\cdot\operatorname{diam} C_j < 1 + \operatorname{diam} C_j$. Let $C_j^k = C_j \times I_j^k$. Then
$A \times [0,1] \subset \bigcup_j C_j \times [0,1] \subset \bigcup_{j,k} C_j \times I_j^k = \bigcup_{j,k} C_j^k$, $\operatorname{diam} C_j^k < 2\cdot\operatorname{diam} C_j \le 2\delta$,

and $\displaystyle\sum_{j=1}^{\infty} \sum_{k=0}^{k(j)} \alpha(s+1) \left[\frac{\operatorname{diam} C_j}{2}\right]^{s+1} < \alpha(s+1) \sum_{j=1}^{\infty} \left(1 + k(j)\right) \left(\operatorname{diam} C_j\right)^{s+1}$

$$\le \frac{\alpha(s+1)}{\alpha(s)} \sum_{j=1}^{\infty} 2\alpha(s) \left[\frac{\operatorname{diam} C_j}{2}\right]^s 2^s \le \frac{\alpha(s+1)}{\alpha(s)} 2^{s+1}\varepsilon .$$

Thus $H_{2\delta}^{s+1}\left(A \times [0,1]\right) = 0$. Since $H_{2\delta}^{s+1}$ is countably subadditive, $H_{2\delta}^{s+1}\left(A \times \bigcup_k [0,k]\right) = 0$. This provides the desired result.

Apply the result three times to obtain the implication that if $H^2\left(Z \cap \partial A\right) = 0$ then $H^5\left(Z \cap \partial A \times \mathbb{R}^3\right) = 0$, hence $H^5\left(Z \cap \partial A \times \mathbb{R}^3-\{0\}\right) = 0$. But $\mathbb{R}^3 \times \left(\mathbb{R}^3-\{0\}\right)$ is the countable union of compact sets and $\psi\circ p$ is locally Lipshitz. Let μ be a coordinate map associated with the open set V in the 5-dimensional manifold $S\mathbb{R}^3$ (i.e., $\mu: V \longrightarrow W \subset \mathbb{R}^5$, W open). Then $\mu\circ\psi\circ p$ is locally Lipshitz, and by subadditivity and the standard result for Lipshitz maps

$$H^5\left(\mu\circ\psi\circ p\left(Z \cap \partial A \times \mathbb{R}^3-\{0\}\right)\right) \le H^5\left(Z \cap \partial A \times \mathbb{R}^3-\{0\}\right) = 0 .$$

Since $H^5 = L^5$ on $\mathbb{R}^5$, where L^5 is the usual Lebesgue measure, it is established that $\psi\left(S_Z\mathbb{R}^3 \times \mathbb{R}_+\right)$ has measure zero in the usual sense of

measure zero for smooth manifolds.

<u>CLAIM 2.</u> Let ψ_∂ be the restriction of ψ to the smooth 5-manifold $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$. Let CP denote the critical points of ψ_∂. Then $\Phi^{-1}(CP)$ has measure zero.

Proof of Claim 2. By Obs.3 $\Phi^{-1}(CP) \subset \psi_\partial(CP)$ and $\psi_\partial(CP)$ is the set of critical values of ψ_∂. By Sard's theorem (Sternberg 1983, Ch2 §3) $\psi_\partial(CP)$ has measure zero.

<u>CLAIM 3.</u> $CP = S(\partial A - Z) \times \mathbb{R}$, where $S(\partial A - Z) \times \mathbb{R}$ is identified with its inclusion in $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$.

Proof of Claim 3. Use the natural coordinates to identify $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$ with $(\partial A - Z) \times S^2 \times \mathbb{R}$ and $S\mathbb{R}^3$ with $\mathbb{R}^3 \times S^2$. Let $(\alpha, x, d): I \longrightarrow (\partial A - Z) \times S^2 \times \mathbb{R}$ be a curve, hence, by the coordinate representation of geodesics in $\mathbb{R}^3$, $\psi\big(\alpha(t), x(t), d(t)\big)$ is identified with $\big(\alpha(t) + d(t)x(t)\ ,\ x(t)\big)$, where $x(t)$ is a unit vector in $\mathbb{R}^3$ for all $t \in I$. The derivative of this curve vanishes if and only if $\dot{x}(0) = 0$ and $\dot{\alpha}(0) + \dot{d}(0)x(0) = 0$. Hence $x(0)$ represents a vector tangent to $\partial A - Z$ and $CP \subset S(\partial A - Z) \times \mathbb{R}$.

Conversely, if $(p, x, d) \in S(\partial A - Z) \times \mathbb{R}$, where $x = \dot{\alpha}(0)$ for a curve $\alpha: I \longrightarrow \partial A - Z$, $\alpha(0) = p$, then the curve $\big(\alpha(-t), x, t+d\big)$ maps under ψ to $\big(\alpha(-t)+(t+d)x\ ,\ x\big)$, which has vanishing derivative at $t = 0$.

The following result will be needed in the next claim.

Sublemma. Let X and Y be manifolds of the same dimension. Let H be an open subset of X and let $f\colon H \longrightarrow Y$ be one-to-one and continuous. Let $g\colon Y \longrightarrow X$ be smooth with $g\circ f = \mathrm{id}|_H$. Let Y_o be the open subset of Y for which g is regular. Then $f^{-1}(Y_o)$ is open in H and $f\colon f^{-1}(Y_o) \longrightarrow f(f^{-1}(Y_o))$ is a diffeomorphism.

Proof of sublemma. Let $x \in f^{-1}(Y_o)$ and let $f(x) \in V \subset Y_o$, where $g|_V$ is a diffeomorphism of V with $g(V)$. By continuity $f^{-1}(V)$ is open. Since $f(x) \in V$, then $x = g\circ f(x) \in g(V)$, or $f^{-1}(V) \subset g(V)$. Therefore, for all $x \in f^{-1}(V)$, $g|_V\circ f(x) = x = g|_V\circ g|_V^{-1}(x)$, hence $f(x) = g|_V^{-1}(x)$ since $g|_V$ is one-to-one. Consequently, every $x \in f^{-1}(Y_o)$ has a neighborhood on which f coincides with a local diffeomorphism, hence the result.

CLAIM 4. Let RP be the regular points of ψ_∂. Then $\Phi^{-1}(RP)$ is open and $\Phi\colon \Phi^{-1}(RP) \longrightarrow S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}_+$ is continuous.

Proof of Claim 4. Let $\Phi(u_1) = \big(\psi\big(u_1, -\tau(u_1)\big)\ ,\ \tau(u_1)\big)$ be a regular point of ψ_∂ in $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$. Let V be a neighborhood of $\Phi(u_1)$ in $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$ in which ψ_∂ is a local diffeomorphism with $\psi_\partial(V) \subset S\mathbb{R}^3$. Shrink V so that $\psi_\partial(V) \subset SM$, which is possible since $\psi_\partial(\Phi(u_1)) = u_1 \in SM$ and SM is open. Moreover, $\psi_\partial(V) \subset SM - \tau^{-1}(\infty)$, since, for any $(v,t) \in V \subset S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$, $\tau(\psi_\partial(v,t)) \le t$.

Summarizing, if $u_1 \in \Phi^{-1}(RP)$, then u_1 is in some open set $\psi_\partial(V) \subset SM - \tau^{-1}(\infty)$, and ψ_∂ is a diffeomorphism on V. It is not necessarily the case for all $(v,t) \in V$ that $\tau(\psi_\partial(v,t)) = t$. However,

it is claimed that there exists a neighborhood $\tilde{V}$ about $\Phi(u_1)$ in V for which this does hold. This suffices to prove the claim, for then on this neighborhood $\Phi(\psi_\partial(v,t)) = \left[\psi\left(\psi_\partial(v,t)\right), -\tau(\psi_\partial(v,t))\right], \tau(\psi_\partial(v,t))\right]$ $= (v,t)$, that is, $\Phi(\psi_\partial(v,t)) \subset \tilde{V}$, with $\psi_\partial(\tilde{V})$ open, hence $\Phi^{-1}(RP)$ is open and on this set Φ is continuous.

To find $\tilde{V}$ consider the open sets in $(\partial A - Z) \times S^2 \times \mathbb{R}$ of the form

$$V_{2\varepsilon} = \left[B_{2\varepsilon} \cap (\partial A - Z)\right] \times W \times \left[\tau(u_1)-2\varepsilon , \tau(u_1)+2\varepsilon\right] \subset \partial A - Z ,$$

where $B_{2\varepsilon}$ is a ball of radius 2ε centered at $\pi(\Phi(u_1))$. Make the natural identification of $(\partial A - Z) \times S^2 \times \mathbb{R}$ with $S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$ and choose ε sufficiently small so that $V_{2\varepsilon} \subset V$. Define V_ε similarly (replace 2ε by ε in all occurrences). It is claimed that

$$\left\{(v,t) \in V_\varepsilon \mid \tau(\psi_\partial(v,t)) \neq t\right\} \subset \left\{(v,t) \in V_\varepsilon \mid \tau(\psi_\partial(v,t)) \leq t-\varepsilon\right\} . \quad (*)$$

For let $u = \psi_\partial(v,t)$ and $(v,t) \in V_\varepsilon$. By the definitions

$$\left[\psi(u,-t) , t\right] = (v,t) \in S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R},$$

while (defining $v' \in S_{\partial A}\mathbb{R}^3$)

$$\left[\psi(u,-\tau(u)) , \tau(u)\right] = (v',\tau(u)) \in S_{\partial A}\mathbb{R}^3 \times \mathbb{R} ,$$

hence

$$\psi\left[\psi(u,-\tau(u)) , \tau(u)-t\right] = \psi(v',\tau(u)-t) = v . \quad (**)$$

If $\tau(u) \neq t$, then $(v',\tau(u))$ cannot be in $V_{2\varepsilon}$, for otherwise, since ψ_∂ is a diffeomorphism on $V_{2\varepsilon}$, then $\psi_\partial(v,t) = u$ and $\psi_\partial(v',\tau(u)) = u$, hence $(v,t) = (v',\tau(u))$ which is a contradiction.

With $(v,t) \in V_\varepsilon$, $\tau(u) \neq t$, and $(v',\tau(u)) \notin V_{2\varepsilon}$, if $\tau(u) \leq t-\varepsilon$, then we are done. If $\tau(u) > t-\varepsilon$, since $t \in (\tau(u_1)-\varepsilon , \tau(u_1)+\varepsilon)$, then for $(v',\tau(u)) \notin V_{2\varepsilon}$ it is necessary that $v' \notin \left[B_{2\varepsilon} \cap (\partial A - Z)\right] \times W$.

But by (**) $\psi(v' \;,\; \tau(u)-t) = v$. That is, in terms of the identification of $S_{\partial A}\mathbb{R}^3$ with $\partial A \times S^2$, $v' = (\pi(v'),w) \in \partial A \times S^2$ and $v = (\pi(v),w) \in \partial A \times S^2$. Thus $\pi(v) \in B_\varepsilon$ while $\pi(v') \notin B_{2\varepsilon}$. Then, again by (**), $|\tau(u)-t| > \varepsilon$, which is a contradiction. This proves (*).

From (*) it follows that $\left\{(v,t) \in V_\varepsilon \mid \tau(\psi_\partial(v,t)) \neq t\right\}$ is closed, for τ is lower semicontinuous by Lemma 1 and consequently $(v,t) \longmapsto \tau(\psi_\partial(v,t)) - t$ is lower semicontinuous. Since $\Phi(u_1) \notin \left\{(v,t) \in V_\varepsilon \mid \tau(\psi_\partial(v,t)) \neq t\right\}$, there is an open set $\tilde{V}$ about $\Phi(u_1)$ in V_ε such that $\tau(\psi_\partial(v,t)) = t$ on $\tilde{V}$. This proves Claim 4.

To complete the proof of the theorem, use Claim 4 and the Sublemma for the case $f = \Phi$, $H = \Phi^{-1}(\dot{R}P)$, $Y = S_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$, $g = \psi_\partial$, $Y_o = RP$. In the statement of the theorem $\mathcal{G} = \Phi^{-1}(RP)$. ∎

Remarks on the proof: (see Figure 4)

1. Note that $\pi(\psi(u,-\tau(u))$ is a point on the "nearest object" to $\pi(u)$ in the direction corresponding to u. That is, the reason for introducing τ and Φ is to keep track of so called occlusions, the

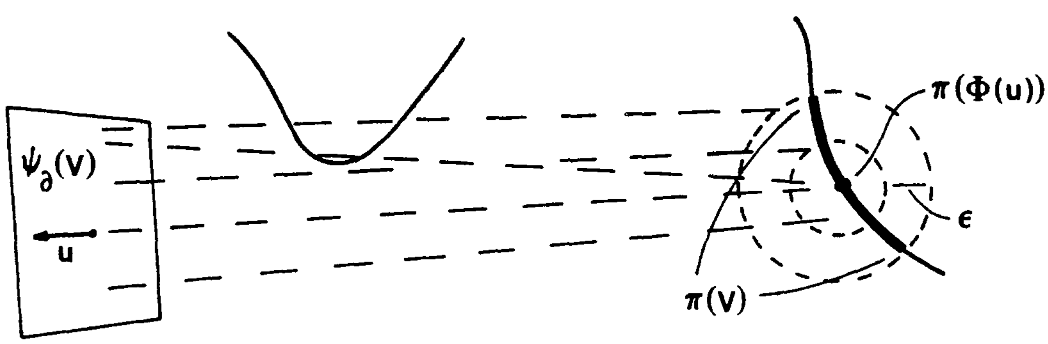


Figure 4

situation in which a geodesic between a point and an object intersects a second object. The map ψ does not respect occlusions: the path $\psi(u,t) = \gamma_u(t)$ "passes through" objects.

2. All of the difficulties in the proof of Claim 4 are associated with the problem of occlusions.

3. The set $\mathcal{Z}$ of measure zero is made up of the union of the "edges" $\Phi^{-1}(S(\partial A - Z) \times \mathbb{R})$ and "corners" $\Phi^{-1}(S_Z\mathbb{R}^3 \times \mathbb{R})$. These edges and corners are subsets of a 5-dimensional manifold. Theorem 10 is a second version of Theorem 8 in which the edges and corners are subsets of a more familiar 2-dimensional manifold that is the analog of an "image plane" in optical devices.

The following corollary describes the manner in which Theorem 8 would typically be used in applied problems.

9. COROLLARY. Let R be a 4-dimensional smooth submanifold of SM such that for every $u \in R$ the curve $\dot{\gamma}_u$ in SM is not tangent to R. Then $R \cap \mathcal{Z}$ has measure zero in the smooth manifold R.

Proof. Recall the flow $\psi: S\mathbb{R}^3 \times \mathbb{R} \longrightarrow S\mathbb{R}^3$, $\psi(u,t) = \dot{\gamma}_u(t)$. Let E denote the associated vector field on SM. By hypothesis E along R is nowhere tangent to R. Then, by the argument in Claim 3 of Theorem 1, ψ restricted to $R \times \mathbb{R}$ is regular. Choose a sufficiently small neighborhood $U = \psi(U \times \{0\}) \subset R$ and a sufficiently small interval

$(-\varepsilon,\varepsilon) \subset \mathbb{R}$ such that $\psi(U \times (-\varepsilon,\varepsilon))$ is an open set of SM. If $\mathcal{Z} \cap U$ does not have measure zero in R, then $\psi(\mathcal{Z} \cap U \times (-\varepsilon,\varepsilon))$ cannot have measure zero in SM (for choose coordinates that map through $U \times (-\varepsilon,\varepsilon)$ and use the product measure properties of Lebesgue measure). From the definition of $\mathcal{Z}$, $\psi(\mathcal{Z} \cap U \times (-\varepsilon,\varepsilon)) = \mathcal{Z} \cap \psi(U \times (-\varepsilon,\varepsilon))$, which must have measure zero by Theorem 8. This is a contradiction. ■

The following theorem is a third version of Theorem 8. It is the analog of Theorem 8 for the case of a single fiber in SM. The notation is as above (e.g., A, M, $\mathcal{G}$, $\mathcal{Z}$). Here $\pi\colon S\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is the projection to the base and $\pi_1\colon S\mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}^3$ is π acting on the first component of the product. This theorem is stated and proved here for completeness. It will not be needed until Chapter 2 of this part. All of the remainder of this chapter depends only on Theorem 8 and its corollaries.

10. THEOREM. Let $A = \bigcup_j A_j$ be non-empty. For every $m \in M$

i. The interior of $\pi^{-1}(m) - \tau^{-1}(\infty)$ is non-empty;

ii. The interior of $\pi^{-1}(m) - \tau^{-1}(\infty)$ contains an open set $\mathcal{G} \cap \pi^{-1}(m)$ on which $\pi_1 \circ \Phi$ is a diffeomorphism of $\mathcal{G} \cap \pi^{-1}(m)$ with its image in ∂A, and $\mathcal{Z} \cap \pi^{-1}(m) = \pi^{-1}(m) - \tau^{-1}(\infty) - \mathcal{G}$ has measure zero as a subset of $\pi^{-1}(m)$;

iii. $\mathcal{Z} \cap \pi^{-1}(m) = \Big[(\pi_1 \circ \Phi)^{-1}(Z) \cup \Phi^{-1}\big[S(\partial A - Z) \times \mathbb{R}\big]\Big] \cap \pi^{-1}(m)$.

Proof. The proof parallels that of Theorem 8. Fix $m \in M$. Let $\vartriangle$ define the section $\vartriangle : \mathbb{R}^3 - \{m\} \longrightarrow S(\mathbb{R}^3 - \{m\}) \times \mathbb{R}$ defined by $\vartriangle(p) = \left[-\psi\left(\dfrac{\exp_m^{-1}(p)}{\|\exp_m^{-1}(p)\|}, \|\exp_m^{-1}(p)\| \right), \|\exp_m^{-1}(p)\| \right]$, where, as usual, $\exp_m : T_m\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is defined by $\exp_m(v) = \gamma_v(1)$. Thus $\psi \circ \vartriangle : \mathbb{R}^3 - \{m\} \longrightarrow \pi^{-1}(m)$ is smooth. (Note $-\psi(u, t) = \psi(-u, -t)$.)

Let Φ_m denote Φ restricted to $\pi^{-1}(m) - \tau^{-1}(\infty) \subset S_m\mathbb{R}^3$. Thus

$$\pi_1 \circ \Phi_m(u) = \exp_m(-\tau(u)u) \qquad (*)$$

The three observations analogous to those in the proof of Theorem 8 can now be made.

Obs.1 $\psi \circ \vartriangle$ is a left inverse for $\pi_1 \circ \Phi_m$.

Obs.2 $\pi_1 \circ \Phi_m$ is one-to-one.

Obs.3 For any set $S \subset \mathbb{R}^3 - \{m\}$, $(\pi_1 \circ \Phi_m)^{-1}(S) \subset \psi \circ \vartriangle(S)$.

<u>CLAIM 1.</u> For $Z \subset \partial A \subset \mathbb{R}^3 - \{m\}$ and $H^2(Z) = 0$, H^2 the 2-dimensional Hausdorff measure, then $(\pi_1 \circ \Phi_m)^{-1}(Z)$ has measure zero in the 2-dimensional mainfold $\pi^{-1}(m)$.

Proof of Claim 1. The proof is a restatement of Claim 1 of Theorem 8 with $\psi \circ \vartriangle$ substituted for ψ and without the need to lift Z to $Z \times \mathbb{R}^3$.

<u>CLAIM 2.</u> Let $(\psi \circ \vartriangle)_\partial$ and $\vartriangle_\partial$ be the restriction of $\psi \circ \vartriangle$ and $\vartriangle$, respectively, to the smooth 2-manifold $\partial A - Z$. Let *cp* denote the

critical points of $(\psi\circ\Delta)_\partial = \psi\circ\Delta_\partial = \psi_\partial\circ\Delta_\partial$. Then $(\pi_1\circ\Phi_m)^{-1}(cp)$ has measure zero in the manifold $\pi^{-1}(m)$.

Proof of Claim 2. By Obs.3 $(\pi_1\circ\Phi_m)^{-1}(cp) \subset (\psi\circ\Delta)_\partial(cp)$. Apply Sard's theorem.

CLAIM 3. Let $CP = \mathrm{S}(\partial A - \mathrm{Z}) \times \mathbb{R}$, where $\mathrm{S}(\partial A - \mathrm{Z}) \times \mathbb{R}$ is identified with its inclusion in $\mathrm{S}_{\partial A-Z}\mathbb{R}^3 \times \mathbb{R}$. Then $(\pi_1\circ\Phi_m)^{-1}(cp) = \Phi_m^{-1}(CP)$.

Proof of Claim 3. Recall CP is the set of critical points of ψ_∂ . By construction Δ is a section, hence $d\Delta_\partial \neq 0$. Consequently, if $p \in \partial A - \mathrm{Z}$ is a critical point of $(\psi\circ\Delta)_\partial$, then $\Delta(p)$ is a critical point of ψ_∂ . That is, $\Delta(p) \subset CP$. This proves the first inclusion of the following string of inclusions. The full string proves the claim.

$$\Phi_m^{-1}(\Delta(p)) \underset{(1)}{\subset} \Phi_m^{-1}(CP) \underset{(2)}{=} \left(\Delta\circ\pi_1\circ\Phi_m\right)^{-1}(CP) = \left(\pi_1\circ\Phi_m\right)^{-1}\left(\Delta^{-1}(CP)\right)$$

$$\underset{(3)}{\subset} \left(\pi_1\circ\Phi_m\right)^{-1}(cp) \underset{(4)}{=} \Phi_m^{-1}(\Delta(cp)) \ .$$

To see (2) note that Δ is defined to satisfy, for $u \in \pi^{-1}(m)$, $\Delta\circ\pi\circ\psi(u,-t) = \left(\psi(u,-t)\ ,\ -t\right)$. Consequently

$$\Delta\circ\pi_1\circ\Phi_m = \Phi_m \ . \qquad (*)$$

To see (3) let $p \in \partial A - \mathrm{Z}$ and $p \in \mathrm{Image}(\pi_1\circ\Phi_m)$ such that $\Delta(p) = (v,t) \in \mathrm{S}(\partial A - \mathrm{Z}) \times \mathbb{R}$. Thus, there exists $u \in \pi^{-1}(m)$ such that

$v = \psi(u,-\tau(u)) \in S(\partial A - Z)$, $\pi\big(\psi(u,-\tau(u))\big) = p$. The curve $\alpha(t) = \pi\big(\psi(u,-\tau(u)+t)\big)$ is tangent to $\partial A - Z$ at $t = 0$, $\dot\alpha(0) = v \neq 0$, but $\psi\circ\Delta\circ\alpha(t) = u$. Thus p is a critical point of $(\psi\circ\Delta)_\partial$.

To see (4) apply (*): for any set S, if $\pi_1\circ\Phi_m(u) \in S$, then $\Delta\circ\pi_1\circ\Phi_m(u) = \Phi_m(u) \in \Delta(S)$.

CLAIM 4. Let rp be the regular values of $(\psi\circ\Delta)_\partial$. Then $(\pi_1\circ\Phi_m)^{-1}(rp)$ is open and $\pi_1\circ\Phi_m : (\pi_1\circ\Phi_m)^{-1}(rp) \longrightarrow \partial A - Z$ is continuous.

Proof of Claim 4. (Since this closely follows the proof of Claim 4, Theorem 8, the presentation is condensed.) Note that the image of $(\psi\circ\Delta)_\partial$ is contained in $\pi^{-1}(m) - \tau^{-1}(\infty)$. Fix $u \in \pi^{-1}(m) - \tau^{-1}(\infty)$, with $\pi_1\circ\Phi_m(u)$ a regular value of $(\psi\circ\Delta)_\partial$. Let V be a neighborhood in $\partial A - Z$ about $\pi_1\circ\Phi_m(u)$ such that $(\psi\circ\Delta)_\partial$ is regular on V. It is claimed that there is a neighborhood $\tilde V$ in V, with $\pi_1\circ\Phi_m(u) \in \tilde V$, such that for all $p \in \tilde V$, if $\Delta(p) = (v,t)$, then $\tau\big(\psi\circ\Delta(p)\big) = t$. This suffices, for then $\pi_1\circ\Phi_m\circ\psi\circ\Delta(p) = \pi_1(v,t) = p$, hence $\pi_1\circ\Phi_m\big((\psi\circ\Delta)_\partial(\tilde V)\big) \subset \tilde V$, with $(\psi\circ\Delta)_\partial(\tilde V)$ open by regularity, hence $(\pi_1\circ\Phi_m)^{-1}(rp)$ is open and on this set $\pi_1\circ\Phi_m$ is continuous.

To find $\tilde V$ let $B_{2\varepsilon}$ be the ball in $\mathbb{R}^3$ of radius 2ε centered at $\pi_1\circ\Phi_m(u)$. Choose $\varepsilon>0$ sufficiently small so that $V_{2\varepsilon} = B_{2\varepsilon} \cap (\partial A - Z) \subset V$, and define B_ε and V_ε similarly.

It is claimed that

$$\left\{p \in V_\varepsilon \mid \Delta(p) = (v,t) \ , \ \tau\big(\psi\circ\Delta(p)\big) \neq t\right\}$$
$$\subset \left\{p \in V_\varepsilon \mid \Delta(p) = (v,t) \ , \ \tau\big(\psi\circ\Delta(p)\big) \leq t-\varepsilon\right\} .$$

For let $u = (\psi\circ\Delta)_\partial(p)$ and $p \in V_\varepsilon$. Let $\Delta(p) = (v,t)$. Then

$$\pi_1\big(\psi(u,-t) \ , \ t\big) = \pi_1(v,t) = p \in V_\varepsilon \subset \partial A - Z \ ,$$
$$\pi_1\big(\psi(u,-\tau(u)) \ , \ -\tau(u)\big) = \pi_1(v',\tau(u)) = \pi(v') \in \partial A - Z \ ,$$

and

$$\psi\big(v',\tau(u)-t)\big) = v, \qquad \pi\circ\psi\big(v',\tau(u)-t\big) = \gamma_{v'}(\tau(u)-t) = p \ . \qquad (*)$$

If $\tau(u) \neq t$, then $\pi(v')$ cannot be in $V_{2\varepsilon}$, for otherwise, since $(\psi\circ\Delta)_\partial$ is a diffeomorphism on $V_{2\varepsilon}$, then $\psi\circ\Delta(p) = \psi\circ\Delta(\pi(v')) = u$, hence $p = \pi(v')$, which is a contradiction of $(*)$. But then, also by $(*)$, since $p \in B_\varepsilon$ and $\pi(v') \notin B_{2\varepsilon}$, $|\tau(u) - t| > \varepsilon$.

By the lower semicontinuity of τ it follows that

$$\left\{ p \in V_\varepsilon \mid \Delta(p) = (v,t) \ , \ \tau(\psi\circ\Delta(p)) \neq t \right\}$$

is closed and does not contain $(\pi_1\circ\Phi_m)(u)$, hence there is an open set about $\pi_1\circ\Phi_m(u)$ in V_ε in which $\tau\big((\psi\circ\Delta)(p)\big) = t$ and $\Delta(p) = (v,t)$. This proves the claim.

The theorem is completed by applying Claim 4 to the sublemma in the proof of Theorem 1. ∎

Why have we bothered with Theorems 8 and 10? Two of the payoffs can now be described. For brevity we shall refer to the situation described in Theorem 8 as the $(S\mathbb{R}^3, SM)$ case, and to that in Theorem 10 as the $(\mathbb{R}^3, S^2)$ case.

The first payoff is a description of so called "edges." In the engineering literature the notion of edge is rarely explicitly defined but rather is described intuitively. Frequently the notion carries the de facto definition of that which the author's edge-finding algorithm finds. What is generally understood to be an edge could be defined as follows.

11. DEFINITION. Let σ be a submersion from a manifold N onto a manifold P. Let S be a subset of N. The boundary $\partial(\sigma(S))$ of the set $\sigma(S)$ is called the edge, or the set of edge points, of $\sigma(S)$ in P.

In Definition 11 it is to be noted that S is the <u>only</u> set in N. There is not sufficient geometry in the notion of submersion to address the case of one set occluding another. This can be partially resolved by considering a disjoint union of such S and the corresponding boundaries. However, the disjoint union includes all boundaries and has no provision for excluding any. (Such an all inclusive projection was used in the proofs of Claims 1 and 2 in Theorem 8.) These considerations are, of course, the motivations behind our definitions of τ and Φ .

Let us find the edges in our two cases. For the $(\mathbb{R}^3, S^2)$ case, $N = \mathbb{R}^3-\{m\}$, $P = \pi^{-1}(m)$, $\sigma = \psi\circ\Delta$, and the set S is a weaker version of our definition of the union of all objects A. The additional structure in our definition of A provides additional information regarding

$\partial(\sigma(S))$ in the case $S = A$. In particular, for the $(\mathbb{R}^3, S^2)$ case, it follows from Theorem 10 that

$$\partial(\sigma(S)) = \partial(\psi\circ\Delta(A)) \subset \Phi^{-1}\left(\pi^{-1}(Z) \cup S(\partial A - Z) \times \mathbb{R}_+\right) . \qquad \text{(edg1)}$$

To see (edg1) it is necessary of first observe that not only $(\pi_1\circ\Phi_m)^{-1}(A) \subset \psi\circ\Delta(A)$ but moreover $(\pi_1\circ\Phi_m)^{-1}(A) = \psi\circ\Delta(A)$ because A is the <u>only</u> set in $\mathbb{R}^3-\{m\}$: for every $p \in A$ $(\pi_1\circ\Phi_m\circ\psi\circ\Delta)(p)$ must be a point in A by the definition of Φ. With this, by Claim 4 of Theorem 10, $(\pi_1\circ\Phi_m)^{-1}(rp)$ is an open set of $\psi\circ\Delta(A)$, hence is in the interior of $\psi\circ\Delta(A)$. Consequently, $(\pi_1\circ\Phi_m)^{-1}(A-rp)$ must contain the boundary $\partial(\psi\circ\Delta(A))$.

The same holds for the $(S\mathbb{R}^3, SM)$ case, where $N = S\mathbb{R}^3 \times \mathbb{R}_+$, $P = SM$ (or a neighborhood in SM or a suitable 4-dimensional submanifold), $\sigma = \psi|_{\psi^{-1}(SM)}$, and $S = S_{\partial A}\mathbb{R}^3 \times \mathbb{R}_+$. Then, as before, since S is the only set, $\Phi^{-1}(S) = \psi(S) \cap SM$, and

$$\begin{aligned}\partial(\sigma(S)) - \partial P &= \partial\left(\psi(S) \cap SM\right) - \partial P \\ &\subset \Phi^{-1}\left[\left(S(\partial A - Z) \times \mathbb{R}_+\right) \cup \left(S_Z\mathbb{R}^3 \times \mathbb{R}_+\right)\right] . \qquad \text{(edg2)}\end{aligned}$$

From the inclusions in (edg1) and (edg2) we can now conclude that for either of our cases the set of edge points is a set of measure zero. We remark that the inclusions are not equalties because our objects may have "corners" (points of non-differentiability) and critical points which may or may not project to edge points.

As is well known, a massive amount of effort has gone into the development of algorithms to find the set $\partial(\sigma(S))$ in P. From our point

of view here, finding $\partial(\sigma(S))$ is a difficult task since $\partial(\sigma(S))$ has measure zero: almost surely, in the sense of probability measure, any point chosen at random is not an edge point. In fact, almost surely every point is a nice point in the sense that it has a neighborhood about it which is, say, diffeomorphically related to a neighborhood in some bundle over ∂A. That is, the nice points constitute an open, dense subset, and the complement of this subset has measure zero. These properties form the basis for the application of geometric method in the remainder of this work.

The second payoff is not that Φ and Φ_m are diffeomorphisms almost everywhere but rather that their compositions with projections are submersions. Thus these compositions injectively pull back differential forms. Throughout the remainder of this chapter we will be considering only the $(S\mathbb{R}^3, SM)$ case. We will return to the $(\mathbb{R}^3, S^2)$ case in the next chapter.

Let p_1 denote the projection $p_1 : S\mathbb{R}^3 \times \mathbb{R} \longrightarrow S\mathbb{R}^3$, and let $\Psi = p_1 \circ \Phi : SM - \tau^{-1}(\infty) \longrightarrow S_{\partial A}\mathbb{R}^3$. The following is immediate.

12. COROLLARY (TO THEOREM 8). There exist an open set $\mathcal{G}$ contained in the interior of $SM - \tau^{-1}(\infty)$ such that $SM - \tau^{-1}(\infty) - \mathcal{G}$ has measure zero as a subset of the smooth manifold SM, and on $\mathcal{G}$ the map Ψ is a submersion.

With this result we can now use differential forms to describe some standard radiometric notions in optics. We first describe *sterance.* In elementary radiometry *sterance* is defined to be "the radiant power emitted from, transmitted through, or reflected off a surface per unit projected area of that surface per solid angle" (Meyer-Arendt 1984, 383). We shall describe a differential form that corresponds to sterance as well as describe a suitable theory of integration.

First, $\partial A - Z$ is orientable, for, by Definition 2 each connected component of $\partial A - Z$ is an open set in the boundary of only one of the objects A_j ($A = \bigcup_j A_j$) , and each ∂A_j by Definition 1 is the boundary of an orientable manifold with boundary. Also, $S_{\partial A-Z}\mathbb{R}^3$ is orientable, for S^2 is orientable and $S_{\partial A-Z}\mathbb{R}^3$ is diffeomorphic to $\partial A - Z \times S^2$.

In describing integration we must use some care since $\partial A - Z$ is not compact. Let ν denote a (global) volume element for $S_{\partial A-Z}\mathbb{R}^3$ (O'Neill 1983, 195). That is, ν is a smooth 4-form on $S_{\partial A-Z}\mathbb{R}^3$ such that, for any orthonormal basis (e_1, e_2, e_3, e_4) of $T_v(S_{\partial A-Z}\mathbb{R}^3)$, $\nu(e_1, e_2, e_3, e_4) = \pm 1$. As usual, any basis (b_1, b_2, b_3, b_4) is said to be positively oriented if $\nu(b_1, b_2, b_3, b_4) > 0$. A coordinate map h of a coordinate chart (U, h), $U \subset S_{\partial A-Z}\mathbb{R}^3$, $h = (x^1, x^2, x^3, x^4): U \longrightarrow \mathbb{R}^4$, is positively oriented if, for $\partial_i = \dfrac{\partial}{\partial x^i}$, $i=1,\cdots,4$, $\nu(\partial_1, \partial_2, \partial_3, \partial_4) > 0$. For such a coordinate neighborhood and positively oriented map define $\int_U \nu$ by

$$\int_U \nu = \int_{h(U)} \nu(\partial_1, \partial_2, \partial_3, \partial_4) \circ h^{-1}$$

where the integral on the right is the Lebesgue integral over the open (hence Lebesgue measurable) set $h(U)$ of the smooth (hence measurable), positive function $\nu(\partial_1, \partial_2, \partial_3, \partial_4) \circ h^{-1}$. Hence, the integral is defined, although it may have the value $+\infty$.

For any open set $E \subset S_{\partial A - Z}\mathbb{R}^3$, let φ_i be a partition of unity subordinate to the countable cover of E by coordinate neighborhoods U_i of an atlas $\{U_i, h_i\}$ with positively oriented coordinate maps. For $h_i = (x_i^1, x_i^2, x_i^3, x_i^4)$, let $\partial_{ij} = \dfrac{\partial}{\partial x_i^j}$ and dx denote $dx_1 dx_2 dx_3 dx_4$. We propose to define $\int_E \nu$ by

$$\int_E \nu = \sum_i \int_{h_i(U_i)} (\varphi_i \chi_E) \circ h_i^{-1} \ \nu(\partial_{i1}, \partial_{i2}, \partial_{i3}, \partial_{i4}) \circ h^{-1} \ dx , \quad E \subset S_{\partial A - Z}\mathbb{R}^3.$$

13. LEMMA. For $E \subset S_{\partial A - Z}\mathbb{R}^3$, $\int_E \nu$ is well defined.

Proof. (1) Each term in the sum is finite, for each φ_i is smooth and compactly supported in $h_i(U_i)$ and ν is smooth. Since all terms of the series are positive, all rearrangements either diverge or converge to the same sum. Thus it suffices to consider the case in which for a choice of $\{U_i, h_i\}$ and $\{\varphi_i\}$ the series converges.

(2) Let $\{V_j, g_j\}$ be a second atlas with positively oriented maps $g_j = (y_j^1, y_j^2, y_j^3, y_j^4)$. Let ψ_j be a partition of unity subordinate to the

cover $\{V_j\}$. For brevity let $(\varphi_i\chi_E)\circ h_i^{-1} = \tilde{\varphi}_i$, $(\psi_j\chi_E)\circ g_j^{-1} = \tilde{\psi}_j$, $H_i = \nu\left(\frac{\partial}{\partial x_i^1},\ldots,\frac{\partial}{\partial x_i^4}\right)\circ h_i^{-1}$, and $G_j = \nu\left(\frac{\partial}{\partial y_i^1},\ldots,\frac{\partial}{\partial y_i^4}\right)\circ g_i^{-1}$. Then, since $\sum_j \psi_j = \sum_i \varphi_i = 1$,

$$\sum_i \int_{h_i(U_i)} \tilde{\varphi}_i H_i\,dx = \sum_i \int_{h_i(U_i)} \left[\sum_j \psi_j\circ h_i^{-1}\right] \tilde{\varphi}_i H_i\,dx$$

$$= \sum_{i,j} \int_{h_i(U_i\cap V_j)} \left[\psi_j\circ h_i^{-1}\right]\tilde{\varphi}_i H_i\,dx$$

$$\overset{(*)}{=} \sum_{i,j} \int_{g_j(U_i\cap V_j)} \left[(\psi_j\chi_E)\circ g_j^{-1}\right]\left[\varphi_i\circ g_j^{-1}\right]G_j\,dy$$

$$= \sum_j \int_{g_j(V_j)} \tilde{\psi}_j G_j\,dy\,,$$

where the equality at (*) is due to the rule for the change of variable and the properties of differential forms. ∎

Our definition holds as well if, for an atlas $\{U_i, h_i\}$, E is the countable union $\bigcup_i h_i^{-1}(h_i(E\cap U_i))$ with each $h_i(E\cap U_i)$ Lebesgue measurable. For if $\{V_j, g_j\}$ is a second atlas, then $g_j\circ h_i^{-1}: h_i(U_i\cap V_j) \longrightarrow V_j$ is smooth, hence $g_j(E\cap V_j) = \bigcup_i g_j\circ h_i^{-1}\left[h_i(E\cap U_i\cap V_j)\right]$ is measurable if each $h_i(E\cap U_i\cap V_j) = h_i(E\cap U_i) \cap h_i(U_i\cap V_j)$ is measurable. By second countability it suffices to require that E satisfy the following

definition.

14. DEFINITION. For any set $E \subset S_{\partial A-Z}\mathbb{R}^3$, E is said to be locally (Lebesgue) measurable if for any coordinate chart (U,h) of $S_{\partial A-Z}\mathbb{R}^3$, $h(E \cap U)$ is Lebesgue measurable.

We are interested here in differential forms. Consequently we can define integration of forms only on the smooth manifold $S_{\partial A-Z}\mathbb{R}^3$. We could equally well define integration on $S_{\partial A}\mathbb{R}^3$ by using, say, the product measure $H^2 \times H^2$ of two 2-dimensional Hausdorff measures for $S_{\partial A}\mathbb{R}^3$ identified with $\partial A \times S^2$ as a subset of $\mathbb{R}^3 \times \mathbb{R}^3$. With this latter definition we can neglect $S_Z\mathbb{R}^3$ for it has measure zero. On $S_{\partial A-Z}\mathbb{R}^3$ the measure of a subset E and the integral $\int_E \nu$ coincide by any of the usual arguments that these two definitions of area of subsets coincide on their common domain of definition. We will not need the measure theoretic definition so we merely make the following definition.

15. DEFINITION. For $E \subset S_{\partial A}\mathbb{R}^3$, $E \cap S_{\partial A-Z}\mathbb{R}^3$ locally measurable, we define $\int_E \nu$ by $\int_E \nu = \int_{E \cap S_{\partial A-Z}\mathbb{R}^3} \nu$.

Defintion 15 has the following obvious extension.

16. DEFINITION. Let $E \subset S_{\partial A}\mathbb{R}^3$, $E \cap S_{\partial A-Z}\mathbb{R}^3$ locally measurable. We say a function f defined almost everywhere on $E \cap S_{\partial A-Z}\mathbb{R}^3$ is locally measurable if the inverse images of Borel sets are locally measurable. For f locally measurable and $f \geq 0$ we define $\int_E f\nu$ by $\int_E f\nu = \int_{E \cap S_{\partial A-Z}\mathbb{R}^3} f\nu$.

Note that on the smooth manifold $S_{\partial A-Z}\mathbb{R}^3$ all of the usual properties of Lebegue integration of nonnegative measurable functions hold. For example, sets of measure zero in $S_{\partial A-Z}\mathbb{R}^3$ can be ignored.

The regard for orientation was not essential in the preceeding defintions. We can neglect orientation by proceding exactly as above with only one modification. (This is an adaptation of Sternberg 1983, Ch.2, sec.3, Integration of densities.) Let ω be a differential n-form on an n-dimensional manifold R. Let D be a locally measurable subset of R. Let $\{\varphi_i\}$ be a partition of unity subordinate to a countable cover of D by coordinate neighborhoods U_i of an atlas $\{U_i, h_i\}$. For $h_i = (x_i^1, \ldots, x_i^n)$, $\partial_{ij} = \dfrac{\partial}{\partial x_i^j}$, and $dx_i = dx_i^1 \ldots dx_i^n$., define

$$\int_D |\omega| = \sum_i \int_{h_i(U_i)} (\varphi_i \chi_D) \circ h_i^{-1} \ |\omega(\partial_{i1}, \ldots, \partial_{in}) \circ h_i^{-1}| \ dx_i \ .$$

The proof that $\int_D |\omega|$ is well defined is just a restatement of the proof of Lemma 13.

To integrate if ω is a smooth n-form defined everywhere except on a set Z of measure zero in R we proceed as in Definition 15.

17. DEFINITION. Let D be locally measurable, $D \subset R$. If ω is a smooth n-form on $R - \mathcal{Z}$ with $\mathcal{Z}$ of measure zero, we define $\int_D |\omega|$ by

$$\int_D |\omega| = \int_{D-\mathcal{Z}} |\omega| \ .$$

These preliminaries are sufficient to complete the discussion of sterance. Recall the map $\Psi: SM-\tau^{-1}(\infty) \longrightarrow S_{\partial A}\mathbb{R}^3$, $\Psi = p_1 \circ \Phi$, which according to Corollary 12 is a submersion on $\mathcal{G} = SM-\tau^{-1}(\infty)-\mathcal{Z}$.

18. DEFINITION. By the *sterance* on $\mathcal{G}$ associated with the differential 4-form $f\nu$ on $S_{\partial A-\mathcal{Z}}\mathbb{R}^3$, $f \geq 0$ and ν a volume element, we mean the differential 4-form on $\mathcal{G}$ defined by $\Psi^*(f\nu)$.

Recall that for any four tangent vectors X_1, X_2, X_3, X_4 in $T_u\mathcal{G}$,

$$\Psi^*(f\nu)\left(X_1, X_2, X_3, X_4\right)(u) = f \circ \Psi(u) \ \nu\left(d\Psi_u X_1, d\Psi_u X_2, d\Psi_u X_3, d\Psi_u X_4\right) ,$$

where $d\Psi_u$ is the differential of Ψ at $u \in \mathcal{G}$.

Since Ψ is a submersion and ν is nonvanishing, $\Psi^*\nu$ is a nonvanishing 4-form on $\mathcal{G}$.

The following result is the motivation for our discussion of integration.

19. PROPOSTION. Let R be a smooth 4-dimensional submanifold of SM that satisfies the conditions in Corollary 9. Let D be a locally measurable subset of R with $D \subset SM-\tau^{-1}(\infty)$ and with Ψ one-to-one on D.

Then $D \cap \mathcal{Z}$ has measure zero in R and

$$\int_D |\Psi^*(f\nu)| = \int_{\Psi(D)} f\nu .$$

Proof. By Definition 16 $\int_{\Psi(D)} f\nu = \int_{\Psi(D)\cap S_{\partial A-Z}\mathbb{R}^3} f\nu$. In addition, the unit tangent bundle $S(\partial A\text{-}Z)$ has measure zero in the manifold $S_{\partial A-Z}\mathbb{R}^3$. Hence, $\int_{\Psi(D)} f\nu = \int_{\Psi(D)\cap S_{\partial A-Z}\mathbb{R}^3-S(\partial A-Z)} f\nu$. Since $\mathcal{Z} = \Psi^{-1}\left(S_Z\mathbb{R}^3 \cup S(\partial A\text{-}Z)\right)$ by Theorem 8, it suffices to show $\int_{D-\mathcal{Z}} |\Psi^*(f\nu)| = \int_{\Psi(D-\mathcal{Z})} f\nu$. This follows essentially from the standard change of variable argument, since Ψ is a smooth map on $D\text{-}\mathcal{Z}$. Briefly, the set where $\Psi^*(f\nu) = 0$ on $D\text{-}\mathcal{Z}$ (that is, where the 4-form $\Psi^*(f\nu)$ annihilates the four basis vectors of the tangent space of $D\text{-}\mathcal{Z}$, or, equivalently, where $d\Psi$ does not have full rank on the tangent space of $D\text{-}\mathcal{Z}$) has image under Ψ of measure zero by Sard's theorem. Thus, we can neglect this set and its image in $\Psi(D)$. Then, when choosing a cover of the remaining open set in $D\text{-}\mathcal{Z}$ by neighborhoods U_i of charts $\{U_i, h_i\}$, we may choose U_i and h_i such that $\Psi\circ h_i^{-1}$ is one-to-one, nonsingular, and $\Psi\circ h_i^{-1}(U_i)$ is the coordinate neighborhoods of a positively oriented chart. This has reduced the problem to the change of variables case for $\mathbb{R}^4$. ■

The requirement that the set D satisfy the requirements of Corollary 9 can be dropped. It will be seen later that $\Psi^*(f\nu) = 0$ on the subset of $D\text{-}\mathcal{Z}$ where the requirements fail. Thus the image of this

subset has measure zero in $\Psi(D)$.

The relationship in Proposition 19 is the basis for our elementary model for the physical process of measuring the energy "detected" at the subset D due to some "emission density" $f\nu$ on $S_{\partial A-Z}\mathbb{R}^3$. In this model there is no time variable: all the objects are assumed stationary and $f\nu$ has no time dependence. To signify this situation we shall say that the sterance is *stationary*. However, we are free to move about in SM. The measuring of energy at D corresponds to the physical process of integrating the power at D over a fixed time interval. The measured value is $\int_D |\Psi^*(f\nu)|$. In applications the value $\int_D |\Psi^*(f\nu)|$ is used to estimate the 4-form $\Psi^*(f\nu)$ at some $u \in D$. To estimate $\Psi^*(f\nu)$ on all of the receptor submanifold R there is typically a covering of R by sets D_j, j=1,..., and the values $\int_{D_j} |\Psi^*(f\nu)|$ are used to estimate $\Psi^*(f\nu)$ at some $u_j \in D_j$. In the next section it will become clear why a single measurement suffices to estimate this vector (4-form) in a five dimensional vector space $(\Lambda^4(T^*SM))$.

The requirement to exclude $\tau^{-1}(\infty)$ is a nuisance. To eliminate this requirement we shall use the fact that the compact set consisting of the union of all objects A is contained in a sufficiently large open ball B. The exterior of B, that is, $\mathbb{R}^3-B$, is added to the set of all objects. Then, since $\partial\left(A \cup \mathbb{R}^3-B\right) = \partial A \cup \partial B$, this new boundary is once again compact. Empty space M is then reduced to $M\cap B$. It is easy to

see from the definition of τ that with this augmented set of objects $\tau^{-1}(\infty) \cap M \cap B = \emptyset$, for the geodesics cannot escape B. Finally, we assume that $f\nu$ is redefined so that its domain is increased to $S_{(\partial A - Z)\cup\partial B}\mathbb{R}^3$.

To conclude this section we now formulate the problem that will be the subject of the remainder of this chapter.

Define, for $m \in M$ (i.e., for $m \in M\cap B$),

$$G(m) = \left\{ \gamma_u(-t) \;\middle|\; u \in S_m\mathbb{R}^3 \; , \; t \in [0,\tau(u)) \right\} .$$

That is, $G(m)$ is the subset of M that can be reached from m by geodesics.

<u>PROBLEM.</u> When is $G(m)$ uniquely determined by the germ of the differential form $\Psi^*(f\nu)$ at m ?

That is, if we know $\Psi^*(f\nu)$ throughout a small neighborhood U in SM, $m \in U$, then what can be said about $G(m)$, in particular, what can be said about the boundary of $G(m)$? In this problem statement we have been extremely generous in comparison to what is usually assumed to be known in various depth vision problems. We briefly list what would be given in the three standard methods. In motion stereo with observer motion it would be assumed that $\Psi^*(f\nu)$ is known along $\pi^{-1}(\alpha(I))$, where π is the projection $\pi: SM \longrightarrow M$ amd $\alpha(I)$ is the path of a curve $\alpha: I \longrightarrow M$. In binocular stereo it would be assumed that $\Psi^*(f\nu)$ is known at $\pi^{-1}(m_1)$ and at $\pi^{-1}(m_2)$, where m_1 and m_2 are two distinct points in M. And in accommodation (depth from focus), it would be assumed that a 2-form is known along a path in M, where the 2-form is

obtained by integrating the 4-form over a neighborhood in each fiber.

It is clear, then, that the solutions we find to the problem in which $\Psi^{*}(fv)$ is known over a neighborhood must be solutions to any of the three standard methods.

1.2 NONUNIQUENESS IN GENERAL

The first part of this section is devoted to checking that the preceding definitions and results are adequate to prove several elementary items. With these items it is shown that even if $\Psi^*(f\nu)$ is given on a neighborhood about a point $u \in SM$, then $G(m)$, the image of geodesic paths in M from m, is not uniquely determined. In this section the convention is continued that $\tau^{-1}(\infty) = \varnothing$ and that $M = M \cap B$ for some open ball B.

Since $\partial A - Z$ is a submanifold of $\mathbb{R}^3$, it has a normal bundle which is a subbundle of $T_{\partial A-Z}\mathbb{R}^3$. Since $\partial A - Z$ has codimension one and is orientable, it has a unit normal vector field N. The field N may be viewed as a section of the bundle $S_{\partial A-Z}\mathbb{R}^3$. We shall use the bracket $\langle , \rangle$ to denote the standard Riemannian metric tensor (that is, the standard inner product of $\mathbb{R}^3$ applied to $T_p\mathbb{R}^3$, $p \in \mathbb{R}^3$, identified with $\mathbb{R}^3$ by the natural isomorphism). The corresponding norm is denoted $\|v\| = \langle v, v\rangle^{1/2}$.

20. LEMMA. Let N be a unit normal field on $\partial A - Z$. Then

$$\langle \Psi(u), (N \circ \pi \circ \Psi)(u) \rangle \neq 0 \ , \quad u \in SM - \mathcal{Z} \ .$$

Proof. From Theorem 8 and the definition of $\Psi = p_1 \circ \Phi$ it follows that $\mathcal{Z} \supset \Psi^{-1}(S(\partial A - Z))$. Hence, $SM - \mathcal{Z} \subset \Psi^{-1}\left(S_{\partial A-Z}\mathbb{R}^3 - S(\partial A - Z)\right)$, which implies $\Psi(SM - \mathcal{Z}) \subset S_{\partial A-Z}\mathbb{R}^3 - S(\partial A - Z)$. ∎

The following definitions are standard ones. Let s_1, s_2, s_3 denote the standard coordinate functions of $\mathbb{R}^3$; that is, $s_i(a_1, a_2, a_3) = a_i$. Let μ_0 denote the volume element $\mu_0 = ds_1 \wedge ds_2 \wedge ds_3$ of $\mathbb{R}^3$.

To similarly express a volume element for $S\mathbb{R}^3$ we shall use interior multiplication and a volume element for $T\mathbb{R}^3$. In particular, for $T\mathbb{R}^3$ the so called natural coordinate functions are $x_1, x_2, x_3, y_1, y_2, y_3$, where $x_i = s_i \circ \pi$ and $y_i = ds_i$, i=1,2,3, with π the natural projection of $T\mathbb{R}^3$. A volume form for $T\mathbb{R}^3$ is then $\mu_1 \wedge \mu_2$, where

$$\mu_1 = dx_1 \wedge dx_2 \wedge dx_3 = \pi^* \mu_0 \ , \text{ and } \ \mu_2 = dy_1 \wedge dy_2 \wedge dy_3 \ .$$

A vector in $T(T\mathbb{R}^3)$ is called *vertical* if it lies in the span of $\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, \frac{\partial}{\partial y_3}$, and it is called *horizontal* if it lies in the span of $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}$. Let the metric tensor for $T\mathbb{R}^3$ be defined by

$$\left\langle \frac{\partial}{\partial y_i}, \frac{\partial}{\partial y_j} \right\rangle = \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle = \left\langle \frac{\partial}{\partial s_i}, \frac{\partial}{\partial s_j} \right\rangle = \delta_{ij} \ ,$$

$$\text{and} \qquad \left\langle \frac{\partial}{\partial y_i}, \frac{\partial}{\partial x_j} \right\rangle = 0 \ , \qquad i,j=1,2,3.$$

Then $\mu_1 \wedge \mu_2$ is a volume element.

Let F denote the vertical vector field

$$F(v) = \sum_i y_i(v) \frac{\partial}{\partial y_i}\Big|_v \ , \quad v \in T\mathbb{R}^3 \ .$$

Note that if $\|v\| = 1$, that is, $v \in S\mathbb{R}^3$, then $\|F(v)\| = 1$.

A p-form (as an alternating multilinear map) is said to be vertical (horizontal) if it vanishes whenever any of the p vectors in the argument is horizontal (vertical). Thus dx_1 and μ_1 are horizontal

forms, $dx_1 \wedge dy_1$ is neither horizontal nor vertical.

Let the interior multiplication of a p-form ω by a vector field X be denoted by $X \rfloor \omega$. That is, $X \rfloor \omega(X_1, \ldots, X_{p-1}) = \omega(X, X_1, \ldots, X_{p-1})$. It follows that $F \rfloor \mu_2$ is a vertical 2-form on $T\mathbb{R}^3$. If the inclusion of $S\mathbb{R}^3$ in $T\mathbb{R}^3$ is denoted by $\iota : S\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3$, then $\iota^*(F \rfloor \mu_2)$ is a vertical 2-form on $S\mathbb{R}^3$ and $\iota^* \mu_1$ is a horizontal 3-form on $S\mathbb{R}^3$. It is straightforward, then, that

$$\mu = \iota^* \mu_1 \wedge \iota^* (F \rfloor \mu_2) = \iota^* \left(-F \rfloor (\mu_1 \wedge \mu_2) \right)$$

is a volume element for $S\mathbb{R}^3$.

In a similar fashion the volume element ν for $S_{\partial A - Z}\mathbb{R}^3$ can be expressed in terms of vertical and horizontal forms. As above, the inclusion maps are used to pull back appropriate forms. For example, let θ be the bundle inclusion map

$$\theta : T_{\partial A - Z}\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3 \ ,$$

and similarly (the subscript 1 to distinguish the unit vector bundle)

$$\theta_1 : S_{\partial A - Z}\mathbb{R}^3 \hookrightarrow S\mathbb{R}^3 \ .$$

Let θ^t denote the adjoint of θ (e.g., $(\theta^t dx_i)(X) = dx_i(\theta X)$, $X \in T_{\partial A - Z}\mathbb{R}^3$). Let N be the previously defined unit normal vector field on ∂A - Z. Then $N \rfloor \theta^t \mu_0$ is a volume element for ∂A - Z, where, as before, we suppress the inclusions T(∂A - Z) in $T_{\partial A - Z}\mathbb{R}^3$ and S(∂A - Z) in $S_{\partial A - Z}\mathbb{R}^3$. With π the projection $\pi : S_{\partial A - Z}\mathbb{R}^3 \longrightarrow \partial A - Z$, $\nu_1 = \pi^*(N \rfloor \theta^t \mu_0)$ is a horizontal 2-form on $S_{\partial A - Z}\mathbb{R}^3$. A vertical 2-form is $\theta_1{}^* \iota^* (F \rfloor \mu_2)$. From the definition of N, F, μ_0, and μ_2 it follows that, up to sign,

$$\nu = \nu_1 \wedge \theta_1{}^* \iota^* (F \rfloor \mu_2) \ .$$

There is one additional 4-form that is needed. Let E denote the horizontal vector field on $T\mathbb{R}^3$ defined by

$$E: T\mathbb{R}^3 \longrightarrow T(T\mathbb{R}^3) \ ,$$
$$E(v) = \sum_i y_i(v)\frac{\partial}{\partial x_i}\Big|_v \ .$$

Then $E\rfloor\mu_1$ is a horizontal 2-form on $T\mathbb{R}^3$, $(E\rfloor\mu_1)\wedge(F\rfloor\mu_2)$ is a 4-form, and, with the inclusion $\iota: S\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3$,

$$\sigma = \iota^*\big[(E\rfloor\mu_1)\wedge(F\rfloor\mu_2)\big] \text{ is a 4-form on } S\mathbb{R}^3.$$

The following simple calculation shows $\sigma = \iota^*\big[(E\wedge F)\rfloor(\mu_1\wedge\mu_2)\big]$:

$$E\rfloor\big(\mu_1\wedge(F\rfloor\mu_2)\big) = (E\rfloor\mu_1)\wedge(F\rfloor\mu_2) + (-1)^3\mu_1\wedge\big[E\rfloor(F\rfloor\mu_2)\big] \ ,$$

while the last term vanishes since μ_2 is vertical; by the same reasoning

$$F\rfloor(\mu_1\wedge\mu_2) = 0 + (-1)^3\mu_1\wedge(F\rfloor\mu_2) \ ;$$

hence

$$(E\rfloor\mu_1)\wedge(F\rfloor\mu_2) = -E\rfloor\big[F\rfloor(\mu_1\wedge\mu_2)\big] \ ,$$

and the right hand side is $(E\wedge F)\rfloor(\mu_1\wedge\mu_2)$.

The results we seek for $S\mathbb{R}^3$ depend, of course, on the structure of $T\mathbb{R}^3$. The following result is elementary, but it provides a technical relationship between $S\mathbb{R}^3$ and $T\mathbb{R}^3$ that will be used repeatedly in proofs. Recall the definition of $\psi_t(u) = \psi(u,t) = \dot{\gamma}_u(t)$ for $u \in S\mathbb{R}^3$, $t \in \mathbb{R}$, and where γ_u is the geodesic determined by u .

21. LEMMA.

i. The horizontal vector field $E: T\mathbb{R}^3 \longrightarrow T(T\mathbb{R}^3)$ is complete and the one parameter group of diffeomorphisms ϕ_t, $t \in \mathbb{R}$, of E is the

geodesic phase flow, $\phi_t(v) = \dot{\gamma}_v(t)$, $t \in \mathbb{R}$, $v \in T\mathbb{R}^3$.

ii. Let $\iota: S\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3$ be the inclusion map, hence $E\circ\iota$ is the restriction of E to $S\mathbb{R}^3$. Then there is a vector field E_1 on $S\mathbb{R}^3$ such that $d\iota\, E_1 = E\circ\iota$. The one parameter group of diffeomorphisms of E_1 is ψ_t, that is, ϕ_t restricted to $S\mathbb{R}^3$. In summary, the following diagram commutes.

$$\begin{array}{ccc}
T(S\mathbb{R}^3) & \xrightarrow{d\iota} & T(T\mathbb{R}^3) \\
{\scriptstyle E_1}\uparrow & & \uparrow{\scriptstyle E} \\
S\mathbb{R}^3 & \xrightarrow{\iota} & T\mathbb{R}^3 \\
{\scriptstyle \psi_t}\downarrow & & \downarrow{\scriptstyle \phi_t} \\
S\mathbb{R}^3 & \xrightarrow{\iota} & T\mathbb{R}^3
\end{array}$$

Proof. For fixed $v \in T\mathbb{R}^3$ the images of v and $E(v) \in T(T\mathbb{R}^3)$ under the standard coordinate functions are

$$v \longmapsto (x_1,x_2,x_3,y_1,y_2,y_3)(v) \in \mathbb{R}^3 \times \mathbb{R}^3,$$

$$E(v) \longmapsto \big[(x_1,x_2,x_3,y_1,y_2,y_3)(v)\ ,\ (y_1,y_2,y_3)(v)\ ,\ (0,0,0)\big]$$
$$\in (\mathbb{R}^3\times\mathbb{R}^3)\times(\mathbb{R}^3\times\mathbb{R}^3)\ ,$$

hence $\phi_t(v) \longmapsto (x_1+ty_1, x_2+ty_2, x_3+ty_3, y_1, y_2, y_3)(v)$. On the other hand $\gamma_u(t) \longmapsto (x_1+ty_1, x_2+ty_2, x_3+ty_3)$, hence $\dot{\gamma}_u(t) = \phi_t(v)$. This proves *i.*, while *ii.* follows from the coordinate representations and the fact that $v \in S\mathbb{R}^3$ if and only if $\sum_i y_i(v)^2 = 1$. ∎

With this lemma we can now say something about the 4-form $\sigma = \iota^*((E\wedge F)\rfloor(\mu_1\wedge\mu_2))$ defined everywhere on $S\mathbb{R}^3$ and about the 4-form

$\Psi^*(f\nu)$ (sterance) defined on $SM\text{-}Z$. In the following we do not require that $M = M \cap B$ for some open ball B.

22. PROPOSITION.

i. $\psi_t^* \sigma = \sigma$, $t \in \mathbb{R}$.

ii. Fix $u_1 \in SM$. For every $t \in \left(-\tau(-u_1), \tau(u_1)\right)$

a) there exists a neighborhood U_t about u_1 such that for all $u \in U_t$, $\tau(u) = \infty$ if and only if $\tau(\psi_t(u)) = \infty$;

b) if $\tau(u_1) \neq \infty$ and $u_1 \notin Z$, then there exists a neighborhood U_t about u_1 such that for all $u \in U_t$ $\Psi\left(\psi_t(u)\right) = \Psi(u)$; in particular, for any p-form ν on $S_{\partial A - Z}\mathbb{R}^3$,

$$\psi_t^*\left(\Psi^*\nu\right) = \Psi^*\nu \text{ almost everywhere on } U_t .$$

Proof. From the diagram of Lemma 21

$$\psi_t^*\sigma = \psi_t^* \iota^*((E\wedge F)\rfloor(\mu_1\wedge\mu_2)) = \iota^*\phi_t^*((E\wedge F)\rfloor(\mu_1\wedge\mu_2)) .$$

Let X_1, X_2, X_3, X_4 be vector fields on $T\mathbb{R}^3$. Then

$$\phi_t^*((E\wedge F)\rfloor(\mu_1\wedge\mu_2))(X_1, X_2, X_3, X_4) =$$
$$\mu_1\wedge\mu_2(E, F, d\phi_t X_1, d\phi_t X_2, d\phi_t X_3, d\phi_t X_4).$$

But $d\phi_t E = E$ since ϕ_t is the one parameter group of diffeomorphisms for E. Moreover, from the standard coordinates, the matrix for $d\phi_t$ is

$$[d\phi_t] = \begin{bmatrix} I & tI \\ 0 & I \end{bmatrix} ,$$

where I is the 3×3 identity matrix. Thus $d\phi_t F = tE + F$. Hence $F =$

$d\phi_t F - tE = d\phi_t F - td\phi_t E$, consequently $E\wedge F = d\phi_t(E\wedge F)$. Therefore

$$\mu_1\wedge\mu_2(E,F,d\phi_t X_1,d\phi_t X_2,d\phi_t X_3,d\phi_t X_4)$$

$$= \phi_t^{*}(\mu_1\wedge\mu_2)(E,F,X_1,X_2,X_3,X_4).$$

But $\phi_t^{*}(\mu_1\wedge\mu_2) = \det[d\phi_t]\mu_1\wedge\mu_2 = (\mu_1\wedge\mu_2)$. This proves *i*.

For *ii.*, for $t \in \mathbb{R}$, $u \in S\mathbb{R}^3$, then if $u \in SM$ and if $t \in \left(-\tau(-u),\tau(u)\right)$, it follows from the definition of τ that $\psi_t(u) \in SM$ and

$$\tau(\psi_t(u)) = t + \tau(u) \ . \qquad (*)$$

The result follows since $\tau^{-1}(\infty)$ is open by Corollary 7. If $\tau(u_1) \neq \infty$ and $u_1 \notin \mathcal{Z}$, then u_1 is in the open set $\mathcal{G} = SM-\tau^{-1}(\infty)-\mathcal{Z}$, as is each $\psi_t(u_1)$, $t \in \left(-\tau(-u_1),\tau(u_1)\right)$. Then

$$\Psi(\psi_t(u)) \overset{\text{def}}{=} \psi\left(\psi_t(u) \ , \ -\tau(\psi_t(u))\right)$$

$$\overset{\text{flow}}{=} \psi\left(u \ , \ t-\tau(\psi_t(u))\right) \overset{(*)}{=} \psi(u,-\tau(u)) = \Psi(u) \ .$$

■

The following result completes the story on the relationship between the two 4-forms. Here we readopt the convention that $\tau^{-1}(\infty) = \emptyset$. Since σ is defined on $S\mathbb{R}^3$ while $\Psi^{*}\nu$ is defined only on $SM-\mathcal{Z}$, we use the inclusion $j: SM-\mathcal{Z} \hookrightarrow S\mathbb{R}^3$ to help keep things sorted out.

23. THEOREM. There exists a function $h: SM\text{-}\mathcal{Z} \longrightarrow \mathbb{R}$ such that

$$\Psi^{*}\nu = hj^{*}\sigma \ .$$

In particular, for $u \in SM\text{-}\mathcal{Z}$, $h(u) = \langle\Psi(u)\ ,\ (N\circ\pi\circ\Psi)(u)\rangle^{-1}$, where N is the unit normal field on $\partial A - Z$.

Proof. The following diagram summarizes the notation.

$$\begin{tikzcd}
T(SM\text{-}\mathcal{Z}) \arrow[r, "dj"] & T(S\mathbb{R}^3) \arrow[r, "dt"] & T(T\mathbb{R}^3) \\
SM\text{-}\mathcal{Z} \arrow[u, "E'_1"] \arrow[r, "j"] \arrow[d, "\Psi"] & S\mathbb{R}^3 \arrow[u, "E_1"] \arrow[r, "t"] \arrow[d, "\psi_t"] & T\mathbb{R}^3 \arrow[u, "E"'] \arrow[d, "\phi_t"'] \\
S_{\partial A-Z}\mathbb{R}^3 \arrow[r, "\theta_1"] & S\mathbb{R}^3 \arrow[r, "t"] & T\mathbb{R}^3
\end{tikzcd}$$

The vector field E'_1 on the open submanifold $SM\text{-}\mathcal{Z}$ of $S\mathbb{R}^3$ is induced by the inclusion j. We claim

$$d\Psi(E'_1) = 0 \ .$$

For consider any function f on $S_{\partial A-Z}\mathbb{R}^3$. By Lemma 21, for $u \in SM\text{-}\mathcal{Z}$,

$$\left[d\Psi(E'_1)(f)\right](u) = \left[E'_1(f\circ\Psi)\right](u) = \frac{d}{dt}\bigg|_{t=0}(f\circ\Psi)(\psi_t(u)) \ .$$

But by Proposition 22 $\Psi(\psi_t(u) = \Psi(u)$ for all t in some small neighborhood about $t=0$, hence the claim.

It follows from the claim that $E'_1\rfloor(\Psi^{*}\nu) = 0$. Locally consider any 1-form e^{*} such that $e^{*}(E'_1) \neq 0$. Then $E'_1\rfloor(e^{*}\wedge\Psi^{*}\nu) = e^{*}(E'_1)\Psi^{*}\nu$. Since $\Psi^{*}\nu$ is a nonvanishing 4-form, $e^{*}\wedge\Psi^{*}\nu$ is a nonvanishing 5-form on $SM\text{-}\mathcal{Z}$, hence there exists a nonvanishing function k on $SM\text{-}\mathcal{Z}$ for which $e^{*}\wedge\Psi^{*}\nu = kj^{*}\mu$, where μ is the volume form $t^{*}(F\rfloor(\mu_1\wedge\mu_2))$ on $S\mathbb{R}^3$. Collecting nonvanishing functions, we have that globally there exists

$h: SM\text{-}\mathcal{Z} \longrightarrow \mathbb{R}$ such that $\Psi^*\nu = hE_1'\rfloor j^*\mu$. From the diagram above, from the expression for μ, and from the definition of σ, it follows that

$$\Psi^*\nu = hj^*t^*\big[(E\wedge F)\rfloor(\mu_1\wedge\mu_2)\big] = hj^*\sigma \ .$$

To evaluate h, fix $u_0 \in SM\text{-}\mathcal{Z}$ and let $t_0 = \tau(u_0)$. Choose $U \subset SM\text{-}\mathcal{Z}$ and $p_1(V) \times I \subset S_{\partial A-\mathcal{Z}}\mathbb{R}^3 \times \mathbb{R}$ such that $u_0 \in U$ and on U the map Φ is a diffeomorphism onto $p_1(V) \times I$. Thus $t_0 \in I$. The map ψ'_{t_0} is defined by the following commuting diagram.

$$\begin{tikzcd}
SM\text{-}\mathcal{Z} \arrow[dd, "\Psi"'] & \supset & U \arrow[rrr, "j"] \arrow[dr, shift left, "\Phi"] & & & S\mathbb{R}^3 \\
& & & p_1(V)\times I \arrow[ul, shift left, "\Phi^{-1}"] \arrow[dl] & \supset p_1(V)\times\{t_0\} & \\
S_{\partial A-\mathcal{Z}}\mathbb{R}^3 & \supset & p_1(V) \arrow[uu, "\psi'_{t_0}"'] \arrow[rrr, "\theta_1"] & & & S\mathbb{R}^3 \arrow[uu, "\psi_{t_0}"']
\end{tikzcd}$$

Hence $\Psi\circ\psi'_{t_0} = \mathrm{id}|_{p_1(V)}$, and locally $\big(\psi'_{t_0}\big)^*\Psi^*\nu = \nu$. Consequently,

$$\nu = \big(\psi'_{t_0}\big)^*(hj^*\sigma) = h\circ\psi'_{t_0}\,(\psi_{t_0}\circ\theta_1)^*\sigma =$$

$$h\circ\psi'_{t_0}\ \theta_1^*\psi_{t_0}^*\ \sigma = h\circ\psi'_{t_0}\ \theta_1^*\sigma \ , \qquad (*)$$

where the last equation is by Propostion 22. Since

$$\theta_1{}^*\sigma = \theta_1{}^*t^*(E\rfloor\mu_1) \wedge \theta_1{}^*t^*(F\rfloor\mu_2) \ ,$$

$$\nu = \pi^*(N\rfloor\theta^t\mu_0) \wedge \theta_1{}^*t^*(F\rfloor\mu_2) \ ,$$

it suffice to compare $\theta_1{}^*t^*(E\rfloor\mu_1)$ and $\pi^*(N\rfloor\theta^t\mu_0)$. Let X_1 and X_2 be any vector fields on $S_{\partial A-\mathcal{Z}}\mathbb{R}^3$ such that $v_1 = d\pi X_1$ and $v_2 = d\pi X_2$ are orthonormal at $\Psi(u_0)$. From the diagram and the definitions, at $\Psi(u_0)$

$$\pi^*(N\rfloor\theta^t\mu_0)\big(X_1, X_2\big) = 1 \ .$$

On the other hand, since $\mu_1 = \pi^*\mu_0$ and $d\pi(E)$ = identity,

$$\theta_1^* \iota^* (E \rfloor \mu_1)\big(X_1, X_2\big) = \mu_0(\iota\circ\theta_1, v_1, v_2) \ .$$

(μ_0 is the volume element of $\mathbb{R}^3$, $\iota\circ\theta_1$ is the inclusion $S_{\partial A-Z}\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3$)

Use these last two results to evaluate (*) at $\Psi(u_0)$, noting that $\psi'_{t_0}(\Psi(u_0)) = u_0$, and $1 = h(u_0)\ \mu_0\big(\Psi(u_0), v_1, v_2\big) = h(u_0)\langle\Psi(u_0), N\circ\pi\circ\Psi(u_0)\rangle$. ∎

Hereafter we shall assume that the function f that appears in the 4-form $f\nu$ is a smooth function on $S_{\partial A-Z}\mathbb{R}^3$.

The following theorem is the nonuniqueness result.

24. THEOREM. Let $u_0 \in SM-\mathcal{Z}$, $\pi(u_0) = m$ (π the projection $S\mathbb{R}^3 \longrightarrow \mathbb{R}$), and let b_m be a ball centered at m such that $b_m \subset M$. Let $U = \pi^{-1}(b_m) \cap \big(SM-\mathcal{Z}\big)$. Let $\{\psi_t \mid t \in \mathbb{R}\}$ be the one parameter group of diffeomorphisms of the vector field E_1 on $S\mathbb{R}^3$. Then $\mathcal{E} = \bigcup_{t\in\mathbb{R}} \psi_t(U)$ is open in $S\mathbb{R}^3$ and

i. There exists a smooth function $g:\mathcal{E} \longrightarrow \mathbb{R}$ such that, in terms of the following inclusion maps

$$i_U: U \longrightarrow \mathcal{E}\ , \qquad i_{\mathcal{E}}: \mathcal{E} \longrightarrow S\mathbb{R}^3\ ,$$

$$j_U: U \longrightarrow SM-\mathcal{Z}\ , \qquad j: SM-\mathcal{Z} \longrightarrow S\mathbb{R}^3\ ,$$

$g i_{\mathcal{E}}^* \sigma$ is a differential 4-form on $\mathcal{E}$ for which

$$j_U^* \Psi^* (f\nu) = i_U^* (g i_{\mathcal{E}}^* \sigma)\ ,$$

where $\Psi: SM-\mathcal{Z} \longrightarrow S_{\partial A-Z}\mathbb{R}^3$ is defined for a fixed choice of a set of objects $\{A_j\}$, $A = \bigcup_j A_j$.

ii. Let B_m be any open ball centered at m such that $b_m \subset B_m$. Let ν_B be the volume element of $S_{\partial(\mathbb{R}^3 - B_m)}\mathbb{R}^3$. Let the single set $\mathbb{R}^3 - B_m$ constitute a second choice for a set of objects $\{A'_j\} = \{\mathbb{R}^3 - B_m\}$, $A' = \mathbb{R}^3 - B_m$, $\partial A' = \partial(\mathbb{R}^3 - B_m) = \partial B_m$. Let ν_B be the volume element of $S_{\partial B_m}\mathbb{R}^3$. Let $\Psi_B : SB_m \longrightarrow S_{\partial B_m}\mathbb{R}^3$ be the map for this set of objects that corresponds to Ψ above. Let j_{UB} be the inclusion $U \hookrightarrow SB_m$. Then there exists a smooth function $f_B : S_{\partial B_m}\mathbb{R}^3 \longrightarrow \mathbb{R}$ such that

$$j_{UB}^{\,*}\left(\Psi_B^*(f_B \nu_B)\right) = j_U^{\,*}\Psi^*(f\nu) = i_U^{\,*}(g i_{\mathcal{E}}^{\,*}\sigma) .$$

Proof: By Theorem 23 and since $j \circ j_U = i_{\mathcal{E}} \circ i_U$,

$$j_U^{\,*}\Psi^*(f\nu) = (f \circ \Psi \circ j_U)\ (h \circ j_U)\ (i_{\mathcal{E}} \circ i_U)^* \ \sigma . \qquad (*)$$

Define $g : \mathcal{E} \longrightarrow \mathbb{R}$ by

$$g\big(\psi_t(u)\big) = \big(f \circ \Psi \circ j_U\big)(u)\ \big(h \circ j_U\big)(u) \qquad \text{for all } u \in U .$$

That this definition does not depend on the choice of u and t in $\psi_t(u)$ follows from the facts *a)* if u_1, $u_2 \in U$, $\psi_{t_1}(u_1) = \psi_{t_2}(u_2)$, then $u_1 = \psi_{t_2 - t_1}(u_2)$; *b)* $\Psi \circ \psi_t = \Psi$ as in Propostion 22; *c)* $h \circ \psi_t = h$ by the formula for h in Theorem 23; and $\pi(U)$ is a ball (convex) contained in M. If we restrict g to U we have

$$g \circ i_U = \big(f \circ \Psi \circ j_U\big)\ \big(h \circ j_U\big) ,$$

and this used in $(*)$ proves *i*.

To prove *ii.* first note that Ψ_B is regular on all of SB_m. That is, the set $\mathcal{Z}_B$ for $\mathbb{R}^3 - B_m$ that corresponds to $\mathcal{Z}$ for $\{A_j\}$ is empty, for ∂B_m is smooth everywhere and $\Psi_B^{-1}(S(\partial B_m)) = \varnothing$. Thus $\pi^{-1}(b_m) \subset SB_m$,

hence $U \subset SB_m$ and the inclusion j_{UB} is well defined. Define f_B to be zero on $\Psi_B \circ j_{UB}(Z)$. It follows from Corollary 9 (to Theorem 8) that this set has measure zero in $S_{\partial B_m}\mathbb{R}^3$. Otherwise define

$$f_B \circ \Psi_B \circ j_{UB}(u) = \frac{\left(f \circ \Psi \circ j_U\right)(u)\ \left(h \circ j_U\right)(u)}{\left(h_B \circ j_{UB}\right)(u)} ,$$

where h_B is given by an application of Theorem 23 and $\left(h_B \circ j_{UB}\right)(u) \neq 0$ by Lemma 20.

From *i.* for the object $\{A_j\}$ we have $j_U^*\Psi^*(f\nu) = i_U^*(gi_{\mathcal{E}}^*\sigma)$ and

$$g\left(\psi_t(u)\right) = \left(f \circ \Psi \circ j_U\right)(u)\ \left(h \circ j_U\right)(u) .$$

Repeating this for the object $\mathbb{R}^3 - B_m$ we have $j_{UB}^*\Psi_B^*(f_B\nu_B) = i_U^*(g_B i_{\mathcal{E}}^*\sigma)$ and

$$g_B\left(\psi_t(u)\right) = \left(f_B \circ \Psi_B \circ j_{UB}\right)(u)\ \left(h_B \circ j_{UB}\right)(u) .$$

But this implies $g = g_B$, hence the result. ∎

Theorem 24 says that the objects are, in effect, irrelevant. The 4-form $\Psi^*(f\nu)$ could equally as well have been defined almost everywhere by using the boundary of a sufficiently large ball. The following immediate corollaries summarize more precisely some observations about Theorem 24.

25. COROLLARIES TO THEOREM 24.

(1) A function g defined on $U \subset SM - Z$ such that $g \circ \psi_t$ is differentiable a $t = 0$ and $\left.\frac{d}{dt}\right|_{t=0}\left(g \circ \psi_t\right)(u) = 0$ for all $u \in U$ determines a function g on $\mathcal{E}(U) = \bigcup_{t \in \mathbb{R}} \psi_t(U)$, hence a 4-form on $\mathcal{E}$, $gi_{\mathcal{E}}^*\sigma$.

(2) For g as in *(1)* and for a ball $B \in \mathbb{R}^3$ such that $\pi^{-1}(B) \supset U$, there exists $f_{\mathrm{B}}: \mathrm{S}_{\partial B_{\mathrm{m}}}\mathbb{R}^3 \longrightarrow \mathbb{R}$ such that $j_{UB}{}^*\Psi_{\mathrm{B}}{}^*(f_{\mathrm{B}}\nu_{\mathrm{B}}) = i_U{}^*(g_{\mathrm{B}} i_{\mathcal{E}}{}^*\sigma)$.

(3) The choice of the ball (the extra object) that made $\tau^{-1}(\infty) = \emptyset$ is arbitrary.

(4) Let the subscripts 1 and 2 be used to distinguish quantities associated with two different sets of objects, both with $\tau^{-1}(\infty) = \emptyset$ for convenience. Then, for every U such that

$$\emptyset \neq U \subset \pi^{-1}(b_{\mathrm{m}}) \cap \left(SM_1 - Z_1\right) \cap \left(SM_2 - Z_2\right) ,$$

b_{m} a ball, for every 4-form $f_1\nu_1$ on $\mathrm{S}_{\partial A_1 - Z_1}\mathbb{R}^3$, there exists a 4-form $f_2\nu_2$ on $\mathrm{S}_{\partial A_2 - Z_2}\mathbb{R}^3$ such that

$$j_{U_1}{}^*\Psi_1{}^*(f_1\nu_1) = j_{U_2}{}^*\Psi_2{}^*(f_2\nu_2) = i_U{}^*(g_{\mathrm{B}} i_{\mathcal{E}}{}^*\sigma) .$$

1.3 A Lambertian submersion

In the last section we saw that any given $\Psi^*(f\nu)$ on a neighborhood could arise, up to a set of measure zero, from essentially any set of objects. This result depended essentially on the fact that $gi_{\mathcal{E}}{}^*\sigma$ coincides with the extension to $\mathcal{E}$ of $\Psi^*(f\nu)$ determined by ψ_t, $t \in \mathbb{R}$. The 4-form $\Psi^*(f\nu)$ depends explicitly on an assumed set of objects while $gi_{\mathcal{E}}{}^*\sigma$ does not. In this section we shift our attention to $gi_{\mathcal{E}}{}^*\sigma$.

We begin by assuming we are given an open set $W \subset S\mathbb{R}^3$ that satisfies the following conditions. Let E_1 be the previously defined horizontal field on $S\mathbb{R}^3$ and let $\{\psi_t\}_{t\in\mathbb{R}}$ be the one parameter group of diffeomorphisms of E_1. (Recall $\psi_t(u) = \dot{\gamma}_u(t)$ for $u \in S\mathbb{R}^3$.) We require that W be *convex with respect to* $\{\psi_t\}$: if $u \in W$ and if for some $t' \in \mathbb{R}$ $\psi_{t'}(u) \in W$, then, for all $t \in [0,t']$, $\psi_t(u) \in W$. (We can equivalently require that, for every $u \in W$, $(\dot{\gamma}_u)^{-1}(W)$ is connected.)

Further, we are given a closed set $\mathcal{Z} \subset S\mathbb{R}^3$ of measure zero which satisfies, for $U = W-\mathcal{Z}$,

$$W \cap \left\{\dot{\gamma}_u(t) \mid u \in U \ , \ t \in \mathbb{R}\right\} \subset U \ .$$

In addition we require that for every $m \in \pi(W)$, $\pi^{-1}(m)\cap W\cap\mathcal{Z}$ has measure zero as a subset of the submanifold $\pi^{-1}(m)$.

As discussed in Section 2, the action of ψ_t generates sets for which the features above are preserved.

$$\mathcal{E}(W) = \bigcup_{t\in\mathbb{R}} \psi_t(W) \ ,$$

$$\mathcal{E}(U) = \bigcup_{t\in\mathbb{R}} \psi_t(U) = \mathcal{E} ,$$

$$\mathcal{E}(\mathcal{Z}) = \bigcup_{t\in\mathbb{R}} \psi_t(\mathcal{Z}) .$$

In addition to W, U, and $\mathcal{Z}$ we are given a map $g|_U : U \longrightarrow \mathbb{R}$. That is, $g|_U$ is defined almost everywhere in W. Further, for each t such that ψ_t is a local diffeomorphism from an open subset of U into U, $g|_U \circ \psi_t = \psi_t^{*} g|_U = g|_U$. As in Section 2, $g|_U$ then has a unique extension to $\mathcal{E}(U)$ which is invariant under ψ_t^{*} for all $t \in \mathbb{R}$ and which coincides with $g|_U$ on U. We denote the extension by g.

As in Section 2, with such a U, $\mathcal{Z}$, and g we define a 4-form by

$$i_U^{*}(g i_{\mathcal{E}}^{*}\sigma) \quad , \quad \sigma = \iota^{*}\left[(E\wedge F)\rfloor(\mu_1\wedge\mu_2)\right] ,$$

with the inclusion maps

$$i_U : U \hookrightarrow \mathcal{E} = \mathcal{E}(U) , \qquad i_{\mathcal{E}} : \mathcal{E} \hookrightarrow S\mathbb{R}^3 ,$$
$$\iota : S\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3 .$$

And $g i_{\mathcal{E}}^{*}\sigma$ is a 4-form on $\mathcal{E}$ with $\psi_t^{*}(g i_{\mathcal{E}}^{*}\sigma) = g i_{\mathcal{E}}^{*}\sigma$, $t \in R$.

We assume g is smooth on U.

We can prove that both σ and $g i_{\mathcal{E}}^{*}\sigma$ are closed. We omit the proof for we have no need for the result. This result is the key element in the classical discussions of the energy entering and leaving a neighborhood.

For the next several definitions we assume once again a fixed set of objects. Suppose $i_U^{*}(g i_{\mathcal{E}}^{*}\sigma) = j_U^{*}\Psi^{*}(f\nu)$, with $f\nu$ a 4-form on $S_{\partial A - Z}\mathbb{R}^3$. Let π denote the projection $\pi : S_{\partial A - Z}\mathbb{R}^3 \longrightarrow \partial A - Z$. Further, suppose that for every $p \in \partial A - Z$ the function g is constant on

$$(\pi\circ\Psi)^{-1}(p) = \{u \in SM \mid (\pi\circ\Psi)(u) = p\} = \{u \in SM \mid \Psi(u) \in \pi^{-1}(p)\} .$$

From Section 2 $g(u) = f\circ\Psi(u) \langle\Psi(u) , (N\circ\pi\circ\Psi)(u)\rangle^{-1}$, hence, if g is constant on $(\pi\circ\Psi)^{-1}(p)$, then for all $v \in \pi^{-1}(p)$

$$f(v) = \text{constant} \langle v , N_p\rangle ,$$

N_p the normal vector at $p \in \partial A - Z$. This expression for f is the usual *Lambertian* condition for the function f . In optics it is often referred to as Lambert's cosine law (Meyer-Arendt 1984).

26. DEFINITION. The function $f:S_{\partial A-Z}\mathbb{R}^3 \longrightarrow \mathbb{R}_+$ is *Lambertian* if there exists a function $f_0:\partial A-Z \longrightarrow \mathbb{R}_+$ such that for every $v \in S_{\partial A-Z}\mathbb{R}^3$ $f(v) = f_0(\pi(v)) \langle v , N(\pi(v))\rangle$. The 4-form $f\nu$, ν a volume element for $S_{\partial A-Z}\mathbb{R}^3$, is said to be a *Lambertian form.*

The following propositions are immediate.

27. PROPOSITION. The 4-form $gi_{\mathcal{E}}{}^*\sigma$ arises from a Lambertian form $f\nu$ on $S_{\partial A-Z}\mathbb{R}^3$ if and only if g is constant on $\Psi^{-1}\left(\pi^{-1}(p)\right)$ for every $p \in \partial A-Z$.

28. PROPOSITION. If $gi_{\mathcal{E}}{}^*\sigma$ arises from a Lambertian form $f\nu$ on $S_{\partial A-Z}\mathbb{R}^3$, if $\theta_1:S_{\partial A-Z}\mathbb{R}^3 \hookrightarrow S\mathbb{R}^3$ is the inclusion map, and if V_1 is the vertical projection of $T_vS\mathbb{R}^3$ to its vertical subspace (the subspace tangent to the fiber at v), then $\left(dg\circ V_1\right)_v = 0$ for every $v \in \theta_1\left(S_{\partial A-Z}\mathbb{R}^3\right) \cap \mathcal{E}(U)$. Here we view $\mathcal{E}(U)$ as an open subset of $S\mathbb{R}^3$ so that $dg\circ V_1$ is a section in $T^*(S\mathbb{R}^3)$ over $\mathcal{E}(U)$.

Recall from Theorem 24 that even if $gi_{\mathcal{E}}^{*}\sigma$ arises from a Lambertian form then it is still the case that for any ball B such that $U \subset \pi^{-1}(B)$ there exists a form $f_{B}\nu_{B}$ on $S_{\partial B}\mathbb{R}^3$ with $i_U^{*}(gi_{\mathcal{E}}^{*}\sigma) = j_{UB}^{*}\Psi_{B}^{*}(f_{B}\nu_{B})$, $j_{UB}: U \hookrightarrow SB$. However, for $p = \pi\left(\Psi_{B}(j_{UB}(U))\right)$, g may not be constant on $\left(\Psi_{B}\circ j_{UB}\right)^{-1}(\pi^{-1}(p))$.

From Theorem 24 and the above discussion it follows that on $\mathcal{E} = \mathcal{E}(U)$ it is sufficient to consider the function g. Let θ_1 be the inclusion $\theta_1 : S_{\partial A-Z}\mathbb{R}^3 \hookrightarrow S\mathbb{R}^3$. We know from Theorem 24 that for a function g with the properties above a codimension one submanifold $\theta_1\left(S_{\partial A-Z}\mathbb{R}^3\right) \cap \mathcal{E}$ of $\mathcal{E}$ exists, but not uniquely, such that $i_U^{*}\left(gi_{\mathcal{E}}^{*}\sigma\right) = j_U^{*}\left(\Psi^{*}(f\nu)\right)$. The remainder of this section is devoted to choosing a codimension one submanifold in $\mathcal{E}$.

We shall say that g is *degenerate* on a neighborhood $U \subset S\mathbb{R}^3$ if g is constant on U. If g is degenerate on U, then it is clear that g is degenerate on $\mathcal{E} = \mathcal{E}(U)$. For example, as in Theorem 24, a degenerate g can arise from f_B defined on $S_{\partial B}\mathbb{R}^3$, B a ball, $B \supset \pi(U)$, by

$$f_B(v) = \text{constant} \langle v\ ,\ (N\circ\pi)(v)\rangle\ , \quad v \in S_{\partial B}\mathbb{R}^3\ ,$$

and N a unit normal field for ∂B. In addition, as in the corollaries to Theorem 24 (Corollaries 25), the same degenerate g can arise from any choice for a set of objects with f having the same definition as f_B. Moreover, for all of these choices f is Lambertian. Hence, if g is degenerate on U, then the condition that f be Lambertian is satisfied trivially. Consequently, we shall exclude the case of g

degenerate on U in the remainder of this section.

Although we do not develop it here, we think the degenerate case is extremely important. This case is useful for "filling in" choices of submanifolds: if g is degenerate on U but Ψ is specifed on the boundary of U, then a choice for Ψ (that is, for $S_{\partial A}\mathbb{R}^3$) on U can be based on variational methods such as minimal surfaces.

If g is nondegenerate we shall use the following tools to choose a codimension one submanifold in $\mathcal{E}$.

Let ω be the fundamental form of $T\mathbb{R}^3$. That is, for s_i the standard coordinate functions for $\mathbb{R}^3$, $x_i = s_i \circ \pi$, $y_i = dx_i$, $i=1,2,3$,

$$\omega = \sum_i y_i \, dx_i \ .$$

Let

$$\Omega = d\omega = \sum_i dy_i \wedge dx_i \ .$$

Recall (Sternberg 1983, Ch.3 sec.7) that Ω is a nondegenerate 2-form and consequently provides a one-to-one correspondence $\alpha_X \longleftrightarrow X_\alpha$, $\alpha_X \in T^*(T\mathbb{R}^3)$, $X_\alpha \in T(T\mathbb{R}^3)$, defined by

$$\alpha_X = X_\alpha \rfloor \Omega \ .$$

We define the horizontal and vertical projections by

$$H: T(T\mathbb{R}^3) \longrightarrow T(T\mathbb{R}^3) \qquad V: T(T\mathbb{R}^3) \longrightarrow T(T\mathbb{R}^3)$$

$$H\left(\frac{\partial}{\partial x_i}\right) = \frac{\partial}{\partial x_i} \qquad V\left(\frac{\partial}{\partial x_i}\right) = 0$$

$$H\left(\frac{\partial}{\partial y_i}\right) = 0 \qquad V\left(\frac{\partial}{\partial y_i}\right) = \frac{\partial}{\partial y_i}$$

Note that $H\left(T(T\mathbb{R}^3)\right)$ and $V\left(T(T\mathbb{R}^3)\right)$ are orthogonal complements of $T(T\mathbb{R}^3)$.

As in the last section we use the inclusion map to pull back objects from $T(T\mathbb{R}^3)$ to $T(S\mathbb{R}^3)$. Since the inclusion map $\iota: S\mathbb{R}^3 \longrightarrow T\mathbb{R}^3$ is a one-to-one immersion, the differential $d\iota$ and the projections H and V determine projections H_1, $V_1: T(S\mathbb{R}^3) \longrightarrow T(S\mathbb{R}^3)$ such that

$$d\iota\circ H_1 = H\circ d\iota \quad\text{and}\quad d\iota\circ V_1 = V\circ d\iota \ .$$

By an abuse of notation we do not distinguish between $T\mathbb{R}^3$ and $S\mathbb{R}^3$ with regard to the coordinates $x_i = s_i\circ\pi$.

We shall use the following notation for vector fields. If X is a vector field, if f is a function, and if ϕ is a diffeomorphism, then X_u is a vector at the point u while $X(f)$ is the function defined by X acting of f. In particular $\left(d\phi\circ X\right)_u = \left(d\phi X\right)_u = d\phi_u\left(X_{\phi^{-1}(u)}\right)$, where we use $d\phi\circ X$ interchangeably with $d\phi X$.

29. LEMMA. The map $T(S\mathbb{R}^3) \longrightarrow T^*(S\mathbb{R}^3)$ defined by $Y \longmapsto Y\rfloor\iota^*\Omega$ provides an isomorphism $V_1T(S\mathbb{R}^3) \longleftrightarrow \left\{\lambda \in H_1T^*(S\mathbb{R}^3) \mid \lambda(E_1) = 0\right\}$. In particular, for g defined on $\mathcal{E} = \mathcal{E}(U) \subset S\mathbb{R}^3$ as above there exists a unique vector field $X_{dg\circ H_1}$ on $\mathcal{E}$ such that

$$V_1\left(X_{dg\circ H_1}\right) = X_{dg\circ H_1} \ ,$$

$$H_1\left(X_{dg\circ H_1}\right) = 0 \ ,$$

and

$$dg\circ H_1 = X_{dg\circ H_1}\rfloor\ \iota^*\Omega \ .$$

Proof. Recall $d\iota\, E_1 = E\circ\iota$ and $E = \sum_i y_i \frac{\partial}{\partial x_i}$. If $Y \in V_1 T_u(S\mathbb{R}^3)$, then $0 = \langle d\iota\, Y,\ F\rangle_{\iota u} = \sum_i dy_i(d\iota Y_u)\ y_i(\iota u)$. But $\sum_i dy_i(d\iota Y_u)\ y_i(\iota u) = \sum_i dy_i(d\iota Y_u)\ dx_i(E(\iota u)) = \left(d\iota Y_u \lrcorner\ \Omega\right)(d\iota\ E_1)$. Since Ω is nondegenerate, the image of $V_1T_u(S\mathbb{R}^3)$ under $Y \longmapsto d\iota Y_u \lrcorner \Omega$ is a 2-dimensional subspace of $HT^*_{\iota u}(T\mathbb{R}^3)$ which is contained in the 2-dimensional subspace $\left\{\lambda \in H_1T^*_{\iota u}(T\mathbb{R}^3) \mid \lambda(E_{\iota u}) = 0\right\}$. Hence these two subspaces coincide. Since $d\iota$ is an isomorphism of $H_1T_u(S\mathbb{R}^3)$ to $HT_{\iota u}(T\mathbb{R}^3)$, ι^*H is injective, and the result follows by considering $\iota^*\left(d\iota Y_u \lrcorner\ \Omega\right)$.

The second part follows from $dg\circ H_1(E_1) = dg(E_1) = \left.\frac{d}{dt}\right|_{t=0} g\left(\psi_t(u)\right) = 0$. In particular, $d\iota\left(X_{dg\circ H_1}\right)\Big|_{\iota u} = \sum_i \frac{\partial g}{\partial x_i}(u)\frac{\partial}{\partial y_i}\Big|_{\iota u}$. ∎

If $dg\circ H_1$ does not vanish on a neighborhood, then the vertical vector field $X_{dg\circ H_1}$ does not vanish on a neighborhood. With this vertical field we can test $dg\circ V_1$.

30. LEMMA. $H_1\circ d\psi_t\circ H_1 = d\psi_t\circ H_1$, that is $d\psi_t\circ H_1$ is horizontal; in particular

$$d\psi_t\left(\frac{\partial}{\partial x_i}\Big|_u\right) = \frac{\partial}{\partial x_i}\Big|_{\psi_t(u)} .$$

Proof: For $\iota: S\mathbb{R}^3 \hookrightarrow T\mathbb{R}^3$ the inclusion, apply $d\iota$ to the left hand side of the statement,

$$dt \circ H_1 \circ d\psi_t \circ H_1 = H \circ dt \circ d\psi_t \circ H_1 = H \circ d\phi_t \circ dt \circ H_1 = H \circ d\phi_t \circ H \circ dt ,$$

and to the right hand side,

$$dt \circ d\psi_t \circ H_1 = d\phi_t \circ H \circ dt .$$

From Section 2 the matrix for $d\phi_t$ in standard coordinates is

$$\begin{bmatrix} I & tI \\ 0 & I \end{bmatrix} .$$

Thus, $d\phi_t\left(\frac{\partial}{\partial x_i}\Big|_u\right) = \frac{\partial}{\partial x_i}\Big|_{\phi_t(u)}$, where $\left\{ \frac{\partial}{\partial x_i}\Big|_u \right\}$ is a basis for $H\left(T_u(T\mathbb{R}^3)\right)$. Hence, $d\phi_t \circ H \circ dt = H \circ d\phi_t \circ H \circ dt$, which suffices to prove the result. ■

31. LEMMA. $\left(\psi_t^*(dg \circ H_1)\right) \circ H_1 = dg \circ H_1$. In particular, if $dg \circ H_1 \neq 0$ on an open set U, then $dg \circ H_1 \neq 0$ on $\mathcal{E}(U)$.

Proof. For $u \in \mathcal{E}$

$$\left[\left(\psi_t^*(dg \circ H_1)\right) \circ H_1\right]_u = \left(dg \circ H_1\right)_{\psi_t(u)} \circ \left(d\psi_t \circ H_1\right)_u \overset{(*)}{=} (dg)_{\psi_t(u)} \circ \left(d\psi_t \circ H_1\right)_u$$

$$= \left[d\left(g \circ \psi_t\right) \circ H_1\right]_u = \left(dg \circ H_1\right)_u ,$$

where $(*)$ is by the preceding Lemma. ■

32. LEMMA. $\Omega = \phi_t^* \Omega$.

Proof. For $u \in \mathcal{E}$

$$\left(\phi_t^* \Omega\right)_u = \left(\phi_t^* \sum_i dx_i \wedge dy_i\right)_u = \sum_i d(x_i \circ \phi_t)_u \wedge d(y_i \circ \phi_t)_u .$$

From the matrix for $d\phi_t$ in the standard coordinates,

$$d(x_1 \circ \phi_t)_u = \left(dx_1 + t\,dy_1\right)_u ,$$

$$d(y_1 \circ \phi_t)_u = \left(dy_1\right)_u ,$$

hence the result. ∎

33. LEMMA. $V_1 \circ d\psi_{-t} \circ X_{dg \circ H_1} = X_{(\psi_t^*(dg \circ H_1)) \circ H_1}$. More generally, if α is a horizontal 1-form such that $\alpha(E_1) = 0$, then $V_1 \circ d\psi_{-t} \circ X_\alpha = X_{\left(\psi_t^*(\alpha)\right) \circ H_1}$.

Proof. For brevity let $X = X_{dg \circ H_1}$. Thus, for $u \in \mathcal{E}$

$$\left[\left(\psi_t^*(dg \circ H_1)\right) \circ H_1\right]_u = \left[\left[\psi_t^*\left(X \lrcorner \iota^*\Omega\right)\right] \circ H_1\right]_u$$

$$= \left(X_{\psi_t(u)} \lrcorner (\iota^*\Omega)_{\psi_t(u)}\right) \circ \left(d\psi_t \circ H_1\right)_u$$

$$= \left((d\iota X)_{\iota\psi_t(u)} \lrcorner \Omega_{\iota\psi_t(u)}\right) \circ \left(d\iota \circ d\psi_t \circ H_1\right)_u$$

$$\text{(apply } \iota\psi_t = \phi_t\iota\ ,\ \phi_t^*\Omega = \Omega)$$

$$= \left((d\iota X)_{\phi_t(\iota(u))} \lrcorner \left(\phi_{-t}^*\Omega\right)_{\phi_t(\iota(u))}\right) \circ \left(d\phi_t \circ d\iota \circ H_1\right)_u$$

$$= \left(d\phi_{-t}(d\iota X)_{\phi_t(\iota(u))} \lrcorner\ \Omega_{\iota(u)}\right) \circ \left(d\iota \circ H_1\right)_u$$

$$= \left((d\iota)_u \left(d\psi_{-t}X\right)_u \lrcorner \Omega_{\iota(u)}\right) \circ \left(d\iota \circ H_1\right)_u$$

$$= \left[\left((d\psi_{-t}X) \lrcorner \iota^*\Omega\right) \circ H_1\right]_u = \left[V_1\left(d\psi_{-t}X\right) \lrcorner \iota^*\Omega\right]_u .$$ ∎

34. PROPOSITION. $\left[X_{dg \circ H_1}\right]_{\psi_t(u)} = V_1\left[d\psi_t\left(X_{dg \circ H_1}\right)_u\right]$, $u \in \mathcal{E}$.

Proof. Note that in Lemma 31 t is arbitrary. Hence,

$$\left[V_1 \circ d\psi_t \circ X_{dg\circ H_1}\right]_u = \left[X_{(\psi_{-t}^{\,*}(dg\circ H_1))\circ H_1}\right]_u = \left[X_{dg\circ H_1}\right]_u ,$$

where the first equality is by Lemma 33 and the last equality is by Lemma 31. Replace u by $\psi_t(u)$. ∎

35. LEMMA. $dg\circ H_1 \circ d\psi_t \circ V_1 = tm^{-1}\left(X_{dg\circ H_1}\right)$, where m^{-1} is the isomorphism from $T_u\mathcal{E}$ to $T_u^*\mathcal{E}$ defined by $\left(m^{-1}(Y_u)\right)(X_u) = \langle Y_u, X_u\rangle$ for vectors X_u and Y_u.

Proof. Use the injection dt. $dt\circ H_1\circ d\psi_t\circ V_1 = H\circ d\phi_t\circ V\circ dt$. But $\left(H\circ d\phi_t\right)\left(\frac{\partial}{\partial y_i}\Big|_u\right) = t\,\frac{\partial}{\partial x_i}\Big|_{\phi_t(u)}$, consequently, for any $X_u \in T_u(S\mathbb{R}^3)$,

$$\left(dt\circ H_1\circ d\psi_t\circ V_1\right)(X_u) = \left(H\circ d\phi_t\circ V\right)(dtX_u)$$

$$= H\circ d\phi_t\Big(\sum_i dy_i(dtX_u)\,\frac{\partial}{\partial y_i}\Big|_u\Big) = t\sum_i dy_i(dtX_u)\,\frac{\partial}{\partial x_i}\Big|_{\phi_t(u)} .$$

Thus, $\left(H_1\circ d\psi_t\circ V_1\right)(X_u) = t\sum_i dy_i(dtX_u)\,\frac{\partial}{\partial x_i}\Big|_{\psi_t(u)}$

$$= t\sum_i dy_i(dtX_u)\,d\psi_t\Big(\frac{\partial}{\partial x_i}\Big|_u\Big) ,$$

where the last equation is by Lemma 30. Since $dg\left[d\psi_t\Big(\frac{\partial}{\partial x_i}\Big|_u\Big)\right] = dg\Big(\frac{\partial}{\partial x_i}\Big|_u\Big) = dy_i\left(dtX_{dg\circ H_1}\right)_u$, it follows that

$$dg\left[\left(H_1\circ d\psi_t\circ V_1\right)(X_u)\right] = t\,\langle X_u\,,\,\left(X_{dg\circ H_1}\right)_u\rangle . \quad ∎$$

36. PROPOSITION.

$$dg\left[\left(H_1 \circ d\psi_t \circ V_1\right)\left(X_{dg\circ H_1}\right)_{\psi_s(u)}\right] = t < X_{dg\circ H_1}, X_{dg\circ H_1} >(\psi_s(u))$$

$$= t < X_{dg\circ H_1}, X_{dg\circ H_1} >(u) \quad .$$

Proof. The first equation is by Lemma 35. For the second equation use Lemmas 29 and 30: $< X_{dg\circ H_1}, X_{dg\circ H_1} >(\psi_s(u)) = \sum_i \left[\frac{\partial g}{\partial x_i}(\psi_s(u))\right]^2 = \sum_i \left[\frac{\partial g}{\partial x_i}(u)\right]^2 = < X_{dg\circ H_1}, X_{dg\circ H_1} >(u)$. ■

Combining all of the above Lemmas we obtain our goal which is contained in the next two theorems.

37. THEOREM.

$$\left(dg\ X_{dg\circ H_1}\right)(\psi_t(u)) = \left(dg\ X_{dg\circ H_1}\right)(u) - t< X_{dg\circ H_1}, X_{dg\circ H_1} >(u) \ .$$

Proof. $\left(dg\ X_{dg\circ H_1}\right)(\psi_t(u)) = dg_{\psi_t(u)}\left(X_{dg\circ H_1}\right)_{\psi_t(u)}$

(apply Proposition 34)

$$= dg_{\psi_t(u)}\left(V_1 \circ d\psi_t \circ X_{dg\circ H_1}\right)_u$$

(add and subtract)

$$= dg_{\psi_t(u)}\left(d\psi_t \circ X_{dg\circ H_1}\right)_u - dg_{\psi_t(u)}\left(H_1 \circ d\psi_t \circ X_{dg\circ H_1}\right)_u$$

$(g\circ\psi_t = g$ and Propostion 36)

$$= \left(dg\ X_{dg\circ H_1}\right)(u) - t < X_{dg\circ H_1},\ X_{dg\circ H_1} >(u)\ . \qquad \blacksquare$$

38. THEOREM. On $\mathcal{E} \cap \left\{dg\circ H_1 \neq 0\right\}$ the function $\left(dg\ X_{dg\circ H_1}\right) : \mathcal{E} \longrightarrow \mathbb{R}$ has zero as a regular value, hence

$\mathcal{E} \cap \left\{dg\ X_{dg\circ H_1} = 0\right\}$ is a submanifold of codimension one.

Proof. Let $\psi_s(u) \in \mathcal{E}$.

$$\left[E_1\left(dg\ X_{dg\circ H_1}\right)\right]\left(\psi_s(u)\right) = \frac{d}{dt}\Big|_{t=0}\left[\left(dg\ X_{dg\circ H_1}\right)\left(\psi_{s+t}(u)\right)\right]$$

(apply Theorem 37)

$$= \frac{d}{dt}\Big|_{t=0}\left[\left(dg\ X_{dg\circ H_1}\right)(u) - (s+t) < X_{dg\circ H_1},\ X_{dg\circ H_1} >(u)\right]$$

$$= - < X_{dg\circ H_1},\ X_{dg\circ H_1} >(u)$$

(by Proposition 36)

$$= - < X_{dg\circ H_1},\ X_{dg\circ H_1} >\left(\psi_s(u)\right) \neq 0\ .$$

In particular, this holds for $s = \dfrac{\left(dg\ X_{dg\circ H_1}\right)(u)}{< X_{dg\circ H_1},\ X_{dg\circ H_1} >(u)}$. $\blacksquare$

This theorem is the primary result of this chapter. In the next section we obtain a modification of this result. We conclude this

section with some interpretation of the theorem.

Of course, the idea behind the theorem is that g is given or known on a neighborhood U in $S\mathbb{R}^3$. Note that if g is given on a codimension one submanifold R that satisfies the conditions of Corollary 9, then there is a unique way to extend g to a neighborhood. In either case, dg and $X_{dg\circ H_1}$ are then known on the neighborhood. Since g has a unique extension to $\mathcal{E} = \mathcal{E}(U)$, then dg and $X_{dg\circ H_1}$ are known on $\mathcal{E}$. The theorem tells us that the set in $\mathcal{E}$ for which the function $dg\ X_{dg\circ H_1}$ is zero is a smooth submanifold. To obtain dg and $X_{dg\circ H_1}$ it is necessary that g be known on a neighborhood.

Since $X_{dg\circ H_1}$ is a vertical vector field, the theorem tells us that the set of points in $\mathcal{E} \subset S\mathbb{R}^3$ on which dg annihilates this vertical field is a smooth submanifold. If g arises from a Lambertian form on a set of objects, then, as was discussed in the beginning of this section, dg is horizontal on $S_{\partial A-Z}\mathbb{R}^3 \subset S\mathbb{R}^3$, that is $dg\circ V_1 = 0$. Recall from the definitions of (Φ and) Ψ that $\Psi(U)$ is the subset of $S_{\partial A-Z}\mathbb{R}^3$ which determines g on U. Consequently, if g arises from a Lambertian form, then $\Psi(U) \subset \mathcal{E} \cap \left\{dg\ X_{dg\circ H_1} = 0\right\}$. It is an obvious corollary to the theorems that for fixed u the path $\left\{\psi_t(u) \mid t \in \mathbb{R}\right\}$ intersects $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ at only one point. Hence $\Psi(U) = \mathcal{E} \cap \left\{dg\ X_{dg\circ H_1} = 0\right\}$.

A second issue is the relationship between $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ and the standard depth vision problems. As was briefly discussed in Section 1,

since g is estimated from integrations, it is physically impossible to determine g (similtaneously) on a neighborhood. (Even an approximation requires a sequence of measurements.) For a fixed set of conditions, the set of possible solutions for g given on, say, a nowhere dense set is certainly larger than the set of solutions for g given on a neighborhood. However, the solution $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ is the <u>only</u> one that is consistent throughout a neighborhood U with the assumption that $dg\circ V_1 = 0$ on the solution set.

A third point is one regarding continuity. The only restriction on g was sufficient differentiabilty and $dg\circ H_1 \neq 0$. This certainly suggests that the results here reformulated appropriately into a problem in terms of function spaces and manifolds would constitute a problem that was well posed.

The final point is how the submanifold is positioned in $\mathcal{E} \subset S\mathbb{R}^3$. Certainly, if g arises from a Lambertian form on $S_{\partial A-Z}\mathbb{R}^3$, then, as was seen, the solution $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ is a subset of $S_{\partial A-Z}\mathbb{R}^3$. However, for arbirary g, with only $dg\circ H_1 \neq 0$, <u>there is no guarantee that $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ lies along fibers in $S\mathbb{R}^3$</u>. (For applications the interpretation is that the "apparent surface" changes position in $\mathbb{R}^3$ as the position of the observer changes.) This observation is of some consequence. It says that for an arbitrary g, $dg\circ H_1 \neq 0$, there is not necessarily a surface in $\mathbb{R}^3$ that is compatible with g throughout a neighborhood U. This certainly suggests that problems that are stated soley in terms of surfaces in $\mathbb{R}^3$ may be ill-posed.

These remarks are made precise by the following corollary to Theorem 38.

39. COROLLARY. For π the projection $\pi: S\mathbb{R}^3 \longrightarrow \mathbb{R}^3$, for $m \in \pi(U)$, the submanifold $\mathcal{E}_m = \mathcal{E}(U) \cap \left[\bigcup_{t\in\mathbb{R}} \psi_t\left(\pi^{-1}(m)\right)\right]$ intersects the submanifold $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ transversely in $\mathcal{E}(U)$. Hence $\mathcal{E}_m \cap \left\{dg\ X_{dg\circ H_1} = 0\right\}$ is a 2-dimensional submanifold of $\mathcal{E}_m$.

Proof. Note that, for every $u \in \mathcal{E}_m$, $\left(E_1\right)_u \in T_u\mathcal{E}_m$, whereas by the proof of Theorem 38 $\left(E_1\right)_u \notin T_u\left\{dg\ X_{dg\circ H_1} = 0\right\}$. And $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ is of codimension one. ∎

Thus each point $m \in \pi(U)$ has an associated 2-manifold in $\mathcal{E}(U)$. When we say that $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ does not necessarily lie along fibers we mean that for a choice of $m \in \pi(U)$ it is not necessarily the case that $\left\{dg\ X_{dg\circ H_1} = 0\right\} = S_N\mathbb{R}^3$, where $N = \pi\left[\ \mathcal{E}_m \cap \left\{dg\ X_{dg\circ H_1} = 0\right\}\ \right]$.

However, $\mathcal{E}_m$ is not $\mathbb{R}^3$. Rather, $\mathcal{E}_m - \pi^{-1}(m) = \bigcup_{t\in\mathbb{R}} \psi_t\left(\pi^{-1}(m)\right) - \pi^{-1}(m)$ is a double covering of $\mathbb{R}^3-\{0\}$. (For $x \in \mathbb{R}^3-\{0\}$, the two points $\left(x, \frac{x}{\|x\|}\right)$ and $\left(x, \frac{-x}{\|x\|}\right)$ are in $(\mathbb{R}^3-\{0\})\times S^2$.) The sign of $\left(dg\ X_{dg\circ H_1}\right)(u)$ in Theorem 37 determines whether $\left\{dg\ X_{dg\circ H_1} = 0\right\}$ is "in front of m" of "behind m." (Note that these two cases are analogous to the distinction between real and virtual images in optics.)

1.4 THE SPHERE BUNDLE OVER A CURVE

In this section we push a bit farther the results of the previous section. In the last section we found that the Lambertian condition determined a codimension one submanifold in the sphere bundle. In this section we find that we can reduce the dimension of everything by one.

Let $a: I \longrightarrow \mathbb{R}^3$ be a curve with nonvanishing tangent vector $\dot{a}(s)$ for all $s \in I$. Since all considerations will be local, let I be sufficiently small so that $a(I)$ is a submanifold (a is a one to one immersion and a homeomorphism into). Let $S_{a(I)}\mathbb{R}^3$ denote the restriction of $S\mathbb{R}^3$ to $a(I)$. Let $S(a(I))$ denote the unit tangent bundle of the manifold $a(I)$ and consider $S(a(I))$ as a subset of $S_{a(I)}\mathbb{R}^3$.

We claim that for every $u \in S_{a(I)}\mathbb{R}^3 - S(a(I))$ there exists a neighborhood $V \subset S_{a(I)}\mathbb{R}^3 - S(a(I))$ about u such that $\bigcup_{t\in\mathbb{R}} \psi_t(V)$ is a submanifold of $S\mathbb{R}^3$. To see this, first note that $\bigcup_{t\in\mathbb{R}} \psi_t(V) = \psi(V\times\mathbb{R})$. (Recall $\psi_t(u) = \psi(u,t)$.) From the proof of Theorem 8, $\psi: V\times\mathbb{R} \longrightarrow S\mathbb{R}^3$ is nonsingular if $V \cap S(a(I)) = \emptyset$. In particular, there exists a neighborhood $V\times(-2\varepsilon,2\varepsilon)$ on which ψ is nonsingular and one to one. Shrink V if necessary so that $\pi(V) \subset B_\varepsilon(\pi(u))$, where $\pi: S\mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is the bundle projection and $B_\varepsilon(\pi(u))$ is the ball of radius ε about $\pi(u)$. Then ψ is one to one on $V\times\mathbb{R}$, for if $\psi_{t_1}(u_1) = \psi_{t_2}(u_2)$, with u_1, $u_2 \in V$, then $\psi_0(u_1) = u_1 = \psi_{t_2-t_1}(u_2)$, hence $|t_1-t_2| < 2\varepsilon$, which contradicts the fact that ψ is one to one on $V\times(-2\varepsilon,2\varepsilon)$. Thus ψ is a one to one

immersion on $V\times\mathbb{R}$. It is also an open map since $\psi\colon S\mathbb{R}^3\times\mathbb{R} \longrightarrow S\mathbb{R}^3$ is an open map. Thus $\bigcup_{t\in\mathbb{R}} \psi_t(V)$ is a submanifold.

Recall the definitions of the open sets W and U of $S\mathbb{R}^3$ defined in the beginning of Section 3. In particular W is convex with respect to $\{\psi_t\}$, $\mathcal{Z} = W-U$ has measure zero in $S\mathbb{R}^3$, and U is convex with respect to $\{\psi_t\}$. By the same type of argument as in Corollary 9, $\mathcal{Z} \cap V$ has measure zero in $S_{a(I)}\mathbb{R}^3$. Let us shrink V by a set of measure zero so that $V \subset U = W - \mathcal{Z}$, so that we may assume as in Section 3 that g, which is smooth on U, is thereby smooth on V . Let $\mathcal{E}_1 = \bigcup_{t\in\mathbb{R}} \psi_t(V)$ and define g on $\mathcal{E}_1$ as before.

With these preliminaries we proceed to refine Section 3 to find a submersion on $\mathcal{E}_1$ determined by the Lambertian condition.

40. PROPOSITION. Let Y be a vertical vector field on $\mathcal{E}(U)$ such that $Y = V_1 \circ d\psi_t \circ Y$. Then the function $dg\ Y$ satisfies

$$(dg\ Y)\circ\psi_t = dg\ Y - t\ \langle X_{dg\circ H_1}\ ,\ Y\rangle\ .$$

Proof. For $u \in \mathcal{E}(U)$

$$\big(dg\ Y\big)\big(\psi_t(u)\big) = dg_{\psi_t(u)} Y_{\psi_t(u)} = \big(dg\circ V_1\circ d\psi_t\big)_u\ Y_u$$

(adding and subtracting)

$$= \big(dg\circ d\psi_t\big)_u\ Y_u - \big(dg\circ H_1\circ d\psi_t\big)_u\ Y_u$$

(apply Lemma 35)

$$= dg_u\ Y_u - t\ \langle X_{dg\circ H_1}\ ,\ Y\rangle_u\ . \qquad \blacksquare$$

Let $u \in S_{a(I)}\mathbb{R}^3 - S(a(I))$ and let $u \in V$, where V is open in $S_{a(I)}\mathbb{R}^3 - S(a(I))$ and $\mathcal{E}_1 = \bigcup_{t\in\mathbb{R}} \psi_t(V)$ is a submanifold of $S\mathbb{R}^3$. Let X be the horizontal lift to $T(S\mathbb{R}^3)$ of the tangent vector to the curve a. That is, $X_u = \sum_i dx_i(\dot{a}(s))\frac{\partial}{\partial x_i}\Big|_u$ for every $u \in \pi^{-1}(a(s))$. Let $\mathfrak{m}_1 : T^*(S\mathbb{R}^3) \longrightarrow T(S\mathbb{R}^3)$ be defined by $d\iota\circ\mathfrak{m}_1\circ\iota^* = \mathfrak{m}$ (recall ι^* is surjective), and let $\omega_1 = \omega\circ d\iota$. Recall $d\iota\mathfrak{m}_1(\omega_1) = \mathfrak{m}(\omega) = E = d\iota E_1$. Thus, the 1-form $\mathfrak{m}_1^{-1}\big(X_u - \omega_1(X_u)(E_1)_u\big)$ annihilates E_1. Consequently, there exists a vertical vector field X_β along a such that

$$\mathfrak{m}_1^{-1}\Big[X_u - \omega_1\big(X_u\big)(E_1)_u\Big] = \big(X_\beta \rfloor \iota^*\Omega\big)_u .$$

If $u \in \pi^{-1}(a(s))$, then $X_\beta \in T_u\mathcal{E}_1$, for by definition $\mathcal{V}_1 T_u\mathcal{E}_1 = \mathcal{V}_1 T_u(S\mathbb{R}^3)$. Consequently, $d\psi_t X_\beta \in T_{\psi_t(u)}\mathcal{E}_1$. We claim the vertical projection is tangent to $\mathcal{E}_1$.

41. LEMMA. $\mathcal{V}_1\circ d\psi_t\circ X_\beta \in T_{\psi_t(u)}\mathcal{E}_1$.

Proof. For $\pi(u) = a(s)$, it is easy to check that

$$d\iota(X_\beta)_u = \sum_i \Big[dx_i(\dot{a}(s)) - \big(\sum_j y_j(\iota u)\ dx_j(\dot{a}(s))\big)y_i(\iota u)\Big]\frac{\partial}{\partial y_i}\Big|_{\iota u} .$$

Then $d\iota\circ\mathcal{V}_1\circ d\psi_t\circ(X_\beta)_u$

$$= \sum_i \Big[dx_i(\dot{a}(s)) - \big(\sum_j y_j(\iota u)\ dx_j(\dot{a}(s))\big)y_i(\iota u)\Big]\frac{\partial}{\partial y_i}\Big|_{\phi_t(\iota u)} ,$$

by the matrix for $d\phi_t$. Clearly $\mathcal{V}_1\circ d\psi_{-t}\circ\mathcal{V}_1\circ d\psi_t(X_\beta)_u$ is tangent to $\mathcal{E}_1$

at $u \in \pi^{-1}(a(s))$ since $V_1 T_u \mathcal{E}_1 = V_1 T_u(\mathrm{S}\mathbb{R}^3)$. We claim $\left(H_1 \circ d\psi_{-t} \circ V_1 \circ d\psi_t\right)(X_\beta)_u$ is in the span of $(E_1)_u$ and X_u. This will suffice to prove the lemma. But $H\left[d\phi_{-t}\left(\frac{\partial}{\partial y_1}\Big|_{\phi_t(tu)}\right)\right] = -t \left.\frac{\partial}{\partial x_1}\right|_{tu}$, so that $H\, d\phi_{-t}\left(V \circ d\phi_t\right) dt(X_\beta)_u = dt\left[-t\, \left(X_u - \omega(X_u)(E_1)_u\right)\right]$. ∎

42. THEOREM. Let the curve $a\colon I \longrightarrow \mathrm{S}\mathbb{R}^3$ determine a submanifold $a(I)$ in $\mathbb{R}^3$ and a subbundle $\mathrm{S}_{a(I)}\mathbb{R}^3\text{-}\mathrm{S}(a(I))$ in $\mathrm{S}\mathbb{R}^3$. Let $V \subset \mathrm{S}_{a(I)}\mathbb{R}^3\text{-}\mathrm{S}(a(I))$ be sufficiently small so that $\mathcal{E}_1 = \bigcup_{t\in\mathbb{R}} \psi_t(V)$ is a submanifold of $\mathrm{S}\mathbb{R}^3$. A vertical vector field X_β along V is defined by $\mathrm{m}_1^{-1}\left(X_u - \omega_1(X_u)(E_1)_u\right) = \left(X_\beta \rfloor \iota^*\Omega\right)_u$, $u \in V$, X the horizontal lift of $\dot{a}$. The vertical vector field $\tilde{X}_\beta = V_1 d\psi_t X_\beta$ along $\mathcal{E}_1$ is tangent to $\mathcal{E}_1$. The function $dg\, \tilde{X}_\beta$ on $\mathcal{E}_1$ satisfies

$$\left(dg\, \tilde{X}_\beta\right) \circ \psi_t = dg\, \tilde{X}_\beta - t\, \langle \tilde{X}_\beta\, ,\, X_{dg\circ H_1} \rangle\, .$$

In particular, $\langle X_\beta\, ,\, X_{dg\circ H_1} \rangle = dg\circ H_1(X)$, so that if $dg\circ H_1(X) \neq 0$ on V, then $dg\, \tilde{X}_\beta$ is a submersion and $\left\{dg\, \tilde{X}_\beta = 0\right\}$ is a codimension one submanifold of $\mathcal{E}_1$.

Proof. Only two things are not covered in Proposition 40 and Lemma 41. One is $\langle X_\beta\, ,\, X_{dg\circ H_1} \rangle = dg\circ H_1(X) = dg(X)$. A brief proof is $\langle X_\beta\, ,\, X_{dg\circ H_1} \rangle = \langle \mathrm{m}_1\left(X_\beta \rfloor \iota^*\Omega\right)\, ,\, \mathrm{m}_1\left(X_{dg\circ H_1} \rfloor \iota^*\Omega\right) \rangle = \langle X - \omega_1(X)E_1\, ,\, \mathrm{m}_1(dg\circ H_1) \rangle$. The second item is that in order to apply Proposition 40

we need $V_1 d\psi_t \tilde{X}_\beta = \tilde{X}_\beta$. But this follows from $V_1 \circ d\psi_t \circ V_1 = V_1 \circ d\psi_t$, which follows from the matrix for $d\phi_t$ by essentially a restatement of the proof of Lemma 30. ∎

We can conclude this section by restating everything from the end of the previous section. All of the the observations are still relevant. The relationship between an arbitrary $dg \circ H_1$ known along a curve and $dg \circ H_1$ due to a Lambertian sterance is the same. Secondly, as before, even though the sterance is stationary, the 2-dimensional manifolds in $\mathcal{E}_m$, m being different points along the curve, are not necessarily over the same subsets in $\mathbb{R}^3$. Finally, as before, the solution submanifolds need not be "in front of" the observer.

2 SOLUTIONS AS VECTOR FIELDS AND 1-FORMS

2.1 MOTIVATION

The subject of this chapter can be viewed as a third problem in the sense that the subject of the previous chapter was the first and second problems. In this point of view, the situation in the previous chapter was, first, that the sterance was specified on a neighborhood and then, second, that it was specified along a curve. In this chapter the situation is, roughly speaking, a generalization of the sterance being given along a curve with the curve not being given. We saw in the previous chapter that the sterance plus the tangent vector to the curve uniquely determined a submanifold solution. It is not surprising, then, that in this chapter the problem is the sense in which both this tangent vector and the submanifold solution can be uniquely determined by the sterance.

In this first section we reexamine the results of Chapter 1 to provide the motivation for the contructions and the questions of this chapter. In this reexamination we first want to clarify what we mean

by knowing the sterance but not knowing the domain. Then we will use the solutions from Chapter 1 to determine the consequences of assuming only this partial knowledge.

The subject of this chapter goes by various names in the applications literature: motion parallax, motion stereo, depth from motion, and optical flow. It would, of course, be possible to motivate this chapter by discussing some of the visual phenomena associated with motion parallax. However, for this the reader is referred to the psychological and engineering literature (Collett and Harkness 1982; Marr 1982; Prazdny 1983). Here we shall restrict ourselves to that for which we have reasonably good definitions and structure. Since we have something resembling this in the first chapter, we will stick to that structure in characterizing our third problem.

It still should be kept in mind that the purpose of this section is motivational. In the second section of this chapter the effort for precision is resumed. We begin the reexamination with a two paragraph summary of what we have.

In the previous chapter we first considered the case in which the function g (as described in the beginning of Section 1.3) was known on a neighborhood U of the sphere bundle, and then we considered the case in which g was known on a neighborhood V in $S_{a(I)}\mathbb{R}^3$, where $a: I \longrightarrow \mathbb{R}^3$ is a curve. In both cases there was a natural choice for vertical vector fields X_α and X_β on the manifolds $\mathcal{E} = \mathcal{E}(U)$ (dimension = 5) and $\mathcal{E}_1 = \mathcal{E}(V) = \bigcup_{t\in\mathbb{R}} \psi_t(V)$ (dimension = 4), respectively, so that, for

$dg \circ H_1 \neq 0$, $\left\{dg\ X_\alpha = 0\right\} \subset \mathcal{E}$ and $\left\{dg\ X_\beta = 0\right\} \subset \mathcal{E}_1$ are codimension one submanifolds.

We also took note of the manifold $\mathcal{E}_m = \mathcal{E}(\pi^{-1}(m))$ (dimension = 3), which is an embedding of $S^2 \times \mathbb{R}$ in $S\mathbb{R}^3$ with $S^2 \times \{0\} \longmapsto \pi^{-1}(m)$. We found, for example, that for each $m = a(s) \in \pi(V)$, $s \in I$, $\mathcal{E}_m$ intersects $\left\{dg\ X_\beta = 0\right\}$ transversely so that $\mathcal{E}_m \cap \left\{dg\ X_\beta = 0\right\}$ is a 2-manifold (see Corollary 39). The projection map π embeds $\mathcal{E}_m \cap \left\{dg\ X_\beta = 0\right\}$ as a submanifold. It was noted that in general the 2-manifolds $\pi\left[\mathcal{E}_{a(s)} \cap \left\{dg\ X_\beta = 0\right\}\right]$ are different for different $s \in I$.

This chapter is motivated in part by the question of the relationship between the family of manifolds $\pi\left[\mathcal{E}_{a(s)} \cap \left\{dg\ X_\beta = 0\right\}\right]_{s\in I}$ and the family of functions $\left\{ g|_{\mathcal{E}a(s)} \right\}_{s\in I}$.

In both of the cases in Chapter 1, either the case of a neighborhood or the case of a curve, the domain of g was $\mathcal{E}$ or $\mathcal{E}_1$. We did not consider the possibility that g could have a time dependence. That is, we did not consider $g: \mathcal{E}\times I \longrightarrow \mathbb{R}$, where $I \subset \mathbb{R}$ is a so called time interval. Let us refer to the case of Chapter 1, $g: \mathcal{E} \longrightarrow \mathbb{R}$, as the *stationary* case for g . Thus, as was noted at several points in Chapter 1, it is physically possible to approximate a stationary g on a neighborhood in $S\mathbb{R}^3$ using sequentially measured samples, whereas it is not physically possible to obtain the measured samples simultaneously.

Let us review some observations regarding the stationary case solutions of Chapter 1.

Observation 1. For the first observation consider the case of g known on a neighborhood in $S\mathbb{R}^3$ as in Section 1.3. Let $F_k: T\mathbb{R}^3 \longrightarrow T\mathbb{R}^3$ be defined by $(x_1, x_2, x_3, y_1, y_2, y_3) \longmapsto (kx_1, kx_2, kx_3, y_1, y_2, y_3)$, $k > 0$, in the standard (or natural) coordinates. Let $F_k: S\mathbb{R}^3 \longrightarrow S\mathbb{R}^3$ be defined in the obvious way by restriction. Let $g' = g \circ F_k^{-1}$. It is straightforward that $F_k\left[\left\{dg\ X_{dg \circ H_1} = 0\right\}\right] = \left\{dg' X_{dg' \circ H_1} = 0\right\}$. This example illustrates the role of the domain of g. In other words, if g were known up to homotheties of the domain, then one would have a family of submanifolds related by homotheties.

Observation 2. The preceding observation is more interesting in the case of the curve a (Section 1.4). Again g is stationary. Note that $S_{a(I)}\mathbb{R}^3 = \bigcup_{s \in I} S_{a(s)}\mathbb{R}^3$. Similarly, with $k(x_1, x_2, x_3) = (kx_1, kx_2, kx_3)$, $S_{ka(I)}\mathbb{R}^3 = \bigcup_{s \in I} S_{ka(s)}\mathbb{R}^3$. Since we can measure g sequentially, we can consider the sequence of functions $g|_{S_{a(s)}\mathbb{R}^3} : S_{a(s)}\mathbb{R}^3 \longrightarrow \mathbb{R}$. We then have the sequence of functions $g'|_{S_{ka(s)}\mathbb{R}^3} = \left(g \circ F_k^{-1}\right)|_{S_{ka(s)}\mathbb{R}^3} : S_{ka(s)}\mathbb{R}^3 \longrightarrow \mathbb{R}$. It should be clear, for a fixed s and for $S_{ka(s)}\mathbb{R}^3$ and $S_{a(s)}\mathbb{R}^3$ identified with S^2 by the natural coordinates and parallel translation, that $g|_{S_{a(s)}\mathbb{R}^3}$ and $\left(g \circ F_k^{-1}\right)|_{S_{ka(s)}\mathbb{R}^3}$ are the same functions on S^2. That is, the sequence of functions parameterized on s is the same. It is only the additional knowledge of the two curves a and ka that distinguishes the cases.

There are even more interesting examples if we drop the requirement that g be stationary. For example, for $s \in I$ let F_k^s be

defined as F_k but with $\exp(sk)$ replacing k. Then $g'|_{S_{\exp(sk)a(s)}\mathbb{R}^3} = \left(g\circ(F_k^s)^{-1}\right)|_{S_{\exp(sk)a(s)}\mathbb{R}^3}$ is the same sequence of functions on S^2 as $g|_{S_{a(s)}\mathbb{R}^3}$, where $S_{a(s)}\mathbb{R}^3$ and $S_{\exp(sk)a(s)}\mathbb{R}^3$ are identified with S^2 by parallel translation and the natural coordinates. Here $g'(u,s) = g\circ(F_k^s)^{-1}(u)$ defines a nonstationary function.

Observation 3. A final observation is the recollection of the fact from the previous chapter that even if g is stationary and $a: I \longrightarrow \mathbb{R}^3$ is known, it is still possible to have a nonconstant sequence of 2-manifolds $\pi\left[\mathcal{E}_{a(s)} \cap \left\{dg\ X_\beta = 0\right\}\right]$ in $\mathbb{R}^3$. The interest here is not that the manifolds "move," but rather that <u>there is no canonical way to define a pointwise correspondence of flow</u> associated with this sequence of manifolds. Although the sequence of manifolds is well defined by $\left\{dg\ X_\beta = 0\right\}$, there is not sufficient structure to uniquely define a flow. (See the discussion of correspondence by Blicher (1984).)

Issues related to these three observations above will occupy us in this chapter. In general, we wish to consider $S_{m_0}\mathbb{R}^3$ for some fixed $m_0 \in \mathbb{R}^3$ and define $\mathcal{E}_0 = \bigcup_{t\in\mathbb{R}} \psi_t(S_{m_0}\mathbb{R}^3)$. The parameter s that was previously the parameter for the curve a is now considered to be a parameter for g. That is. we wish to study functions $g: \mathcal{E}_0\times I \longrightarrow \mathbb{R}$, $I \subset \mathbb{R}$, which satisfy $g\left(\psi_t(u)\ ,\ s\right) = g\left(u\ ,\ s\right)$ for $(u,s) \in \mathcal{E}_0\times I$. The coordinate s is to be interpreted as time. As usual, we assume g is smooth where needed. Such a function g on $\mathcal{E}_0\times I$ would arise, for

example, from g' described at the end of Observation 2.

In this formulation we have dropped all information about the path a and the neighborhood U. In the results of Chapter 1 precise information about the path was required and g was to be stationary. This amounts to assuming that $g: \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ arises from a rigid motion of $\mathbb{R}^3$. Here we wish to remove this restriction and to consider a wider class of motions. (For applications we have in mind examples of $g: \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ that arise not only from observer motion but also from moving objects such as water waves, animals, wind blown grain fields, flapping flags, as well as isometries of $\mathbb{R}^3$.) In fact, rather than assume an isometry, we wish to determine to what extent local analysis of $g: \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ can be used to "detect" isometries.

In Observation 2 there is a preview of the type of degeneracies that will be faced. The example involving F_k^S is equivalent to the observation that $g: \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ does not vary if the objects about an observer at m_0 collapse (are retracted to m_0) or expand along radial lines. Thus $g: \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ cannot "detect" such motion; i.e., g is invariant under such motion. From Observation 3 there is the second preview that, even if a moving manifold is specified, an associated flow is not necessarily uniquely determined.

A first result from these remarks is that we may drop the pretense of working in the sphere bundle. With the understanding that we are interested only in points not at the origin, we have the following.

Remark. Without loss of generality we may replace $g\colon \mathcal{E}_0 \times I \longrightarrow \mathbb{R}$ with $g\colon (\mathbb{R}^3-\{0\})\times I \longrightarrow \mathbb{R}$.

Reason. We can identify $\mathbb{R}^3-\{0\}$ with $\bigcup_{t<0} \psi_t(S_{m_0}\mathbb{R}^3) \subset \bigcup_{t\in\mathbb{R}} \psi_t(S_{m_0}\mathbb{R}^3) = \mathcal{E}_0$, and $\mathbb{R}^3-\{0\}$ suffices, for we may reflect, if necessary, any of the codimension one submanifolds of Chapter 1 (e.g., ∂A, $\pi\left[\left\{dg\ X = 0\right\}\right]$) through the origin because of the ψ_t invariance property of g .

The problem of selecting (moving) submanifolds in $(\mathbb{R}^3-\{0\})\times I$ that are consistent with a given $g\colon (\mathbb{R}^3-\{0\})\times I \longrightarrow \mathbb{R}$ has at least the degeneracy of the problems in Chapter 1. In Chapter 1 we had nonuniqueness in general, but unique Lambertian submanifolds. We seek similar conditions for a well posed problem for selecting submanifolds consistent with $g\colon (\mathbb{R}^3-\{0\})\times I \longrightarrow \mathbb{R}$. In this chapter we shall make an assault and some progress on this problem. Our progress will at times consist of solving a subproblem which we call the *still picture problem.*

DEFINITION. We say that a problem involving $g\colon (\mathbb{R}^3-\{0\})\times I \longrightarrow \mathbb{R}$ satisfies the *still picture condition* if whenever g satisfies

$$g(m,s) = g(m,s_0)\ ,\quad (m,s) \in (\mathbb{R}^3-\{0\})\times I\ ,\quad s_0 \text{ fixed in } I,$$

then the only possible flow on $\mathbb{R}^3-\{0\}$ is the identity.

Since, for example, the flow on $(\mathbb{R}^3-\{0\})\times I$ given by $\varphi_s(m,s_0) = \left(e^{-s}m\ ,\ s_0+s\right)$ does not change g, then there obviously must be some

additional structure for the problem to satisfy the still picture condition. In this study we have adopted this still picture condition as our first question to be asked (and answered) in seeking to understand mathematical structures for modeling depth vision phenomena.

2.2 Space, basis vectors, flows, and forms

In the last section $(\mathbb{R}^3-\{0\})\times I$ was identified as the space of interest. In this section definitions and some elementary relationships are presented. As in Chapter 1, we are pursuing local results, and we will ignore any closed set of measure zero. The structure of the ignored sets is contained in Theorem 10 of Chapter 1: we are ignoring, among other things, "edges." Since everything in this chapter is local (i.e., we only need some open set around the point of interest), it suffices to know that the ignored set is closed and of measure zero. The additional structure given in the theorem can be neglected.

Let $(m,t) \in (\mathbb{R}^3-\{0\})\times I$, where I is an open interval in $\mathbb{R}$. In this chapter t will always be an element of I. The natural coordinate functions for $(\mathbb{R}^3-\{0\})\times I$ will be denoted by x_1,x_2,x_3,s : if $m = (m_1,m_2,m_3) \in \mathbb{R}^3-\{0\}$, then $x_i(m,t) = m_i$, $i=1,2,3$, and $s(m,t) = t$. $(\mathbb{R}^3-\{0\})\times I$ is equipped with the standard metric $<\frac{\partial}{\partial x_i} , \frac{\partial}{\partial x_j}> = \delta_{ij}$, $<\frac{\partial}{\partial x_i} , \frac{\partial}{\partial s}> = 0$, $<\frac{\partial}{\partial s} , \frac{\partial}{\partial s}> = 1$. The vector field $\frac{\partial}{\partial s}$ determined by the natural coordinates will often be denoted ∂_s . Equivalently, ∂_s = grad(s).

For $m = (m_1,m_2,m_3)$, let $|m|^2 = \sum_i m_i^2$. The smooth, positive valued function ρ on $(\mathbb{R}^3-\{0\})\times I$ is defined by $\rho(m,t) = |m|$.

1. DEFINITION. The *position vector field* $\mathcal{R}$ on $(\mathbb{R}^3-\{0\})\times I$ is defined by $\mathcal{R} = \sum_i x_i \frac{\partial}{\partial x_i}$. The one parameter group of diffeomorphisms of the vector field $\mathcal{R}$ is denoted by ϑ_r , $r \in \mathbb{R}$.

It is fairly easy to see, for $(m,t) \in (\mathbb{R}^3-\{0\})\times I$, that $\vartheta_r(m,t) = \gamma_{\mathcal{R}(m,t)}(e^r-1)$, where $\gamma_{\mathcal{R}(m,t)}$ is the geodesic determined by the initial tangent vector $\mathcal{R}(m,t)$. In particular, $\mathcal{R}$ is a complete vector field on $(\mathbb{R}^3-\{0\})\times I$; that is, ϑ_r is indeed defined for all $r \in \mathbb{R}$.

The following elementary consequences are noted: $ds\ \mathcal{R} = 0$; $s\circ\vartheta_r = s$; $d\rho\ \mathcal{R} = \rho$; $\ln\rho$ is well defined on $(\mathbb{R}^3-\{0\})\times I$ and $d\ln\rho\ \mathcal{R} = 1$; $\mathcal{R} = \rho\ \mathrm{grad}\rho$. Since $\rho\circ\vartheta_r = e^r\rho$ and since ϑ_r^* commutes with d, $\vartheta_r^*\ d\ln\rho = d\ln\rho$.

In this chapter two types of figures will be used. Each type has its unique ambiguity. In Figure 5 the vectors $\mathcal{R}(m,t)$ and $\partial_s(m,t)$ are illustrated for a fixed $(m,t) \in (\mathbb{R}^3-\{0\})\times I$. In Figure 5a the first factor $\mathbb{R}^3-\{0\}$ of $(\mathbb{R}^3-\{0\})\times I$ is displayed ambiguously as a quadrant of a two dimensional plane whereas the second factor I is displayed unambiguously. In Figure 5b the second factor I is ambiguously displayed. The two dimensional plane in Figure 5b is used to represent either the slice $(\mathbb{R}^3-\{0\})\times\{t\}$ or the image of the projection $\pi\colon (\mathbb{R}^3-\{0\})\times I \longrightarrow \mathbb{R}^3-\{0\}$.

We wish to consider vector fields on $(\mathbb{R}^3-\{0\})\times I$ that correspond to the so called nonautonomous or time dependent vector fields (Arnold 1973, Ch.2 sec.8). We also wish to restrict the consideration to

those vector fields that are consistent with the scaling observations of the previous section.

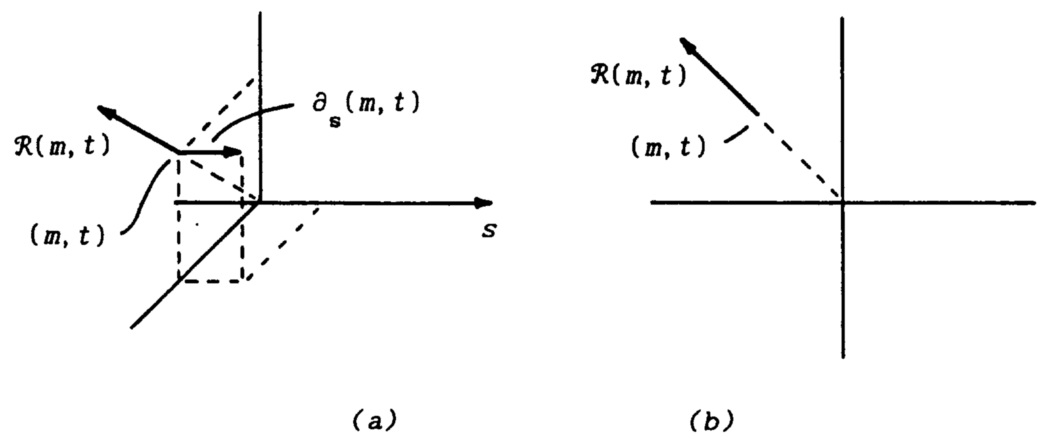


Figure 5

2. DEFINITION. An *admissible vector field* is a vector field X on a neighborhood $U \subset (\mathbb{R}^3-\{0\})\times I$ such that

i. $U = \bigcup_{r\in\mathbb{R}} \vartheta_r(U)$,

ii. $ds\, X = 1$,

iii. $[X,\mathcal{R}] = 0$, where $[\,,]$ is the Lie bracket.

The local one parameter group of X is denoted $\{\varphi_\Delta\}$ (Warner 1971, Definition 1.49). From Definition 2 one has several elementary consequences. First one can apply the standard result that linear

independent vector fields with vanishing Lie bracket have commuting one parameter groups on a sufficiently small neighborhood (Bishop and Goldberg 1968, Theorem 3.7.1). Since $ds\, X = 1$ and $ds\, \mathcal{R} = 0$, X and $\mathcal{R}$ are linearly independent on U. Since $\mathcal{R}$ is complete, there is the following (Figure 6).

3. LEMMA. $d\vartheta_r X = X$ on U. If φ_Δ is defined on $V \subset (\mathbb{R}^3-\{0\})\times I$ for $\Delta \in J \subset \mathbb{R}$, then φ_Δ is defined on $\bigcup_{r\in\mathbb{R}} \vartheta_r(V)$ and $\varphi_\Delta \circ \vartheta_r = \vartheta_r \circ \varphi_\Delta$.

Proof. For any $(m,t) \in U$, cover the compact set $K_n = \bigcup_{-n\le r\le n} \vartheta_r(m,t)$ by a finite collection of neighborhoods in which the local one parameter groups commute, hence in which $d\vartheta_r\, X = X$. By the completeness of R and by the group property of ϑ_r it follows that this holds on K_n , hence on U . It is then easy to check that the integral curve of X at $\vartheta_r(m,t)$ is $(\vartheta_r \circ \varphi_\Delta)(m,t)$, and the last statement follows by uniqueness. ∎

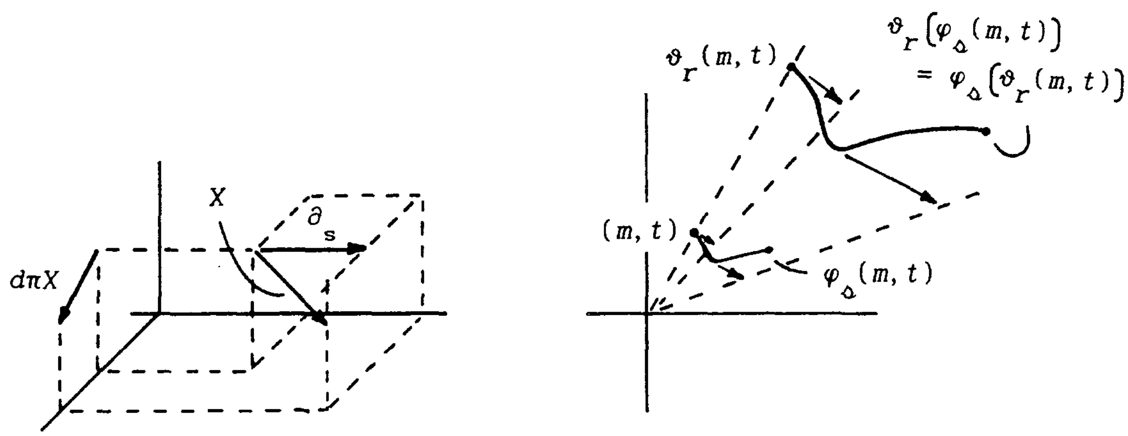


Figure 6

An additional consequence of Definition 2 is the following. Since $ds\, X = 1$, then, for (m,t) in the domain of φ_Δ , $s\big(\varphi_\Delta(m,t)\big) = t + \Delta = s(m,t) + \Delta$. Thus, since φ_Δ^* commutes with d, $\varphi_\Delta^*\, ds = ds$.

We digress briefly for a few remarks on motivation and modeling regarding the choice of definition for the admissible vector field. First, the flow φ_Δ determined by a time varying vector field X is the obvious generalization of the transformations discussed in the Observations of Section 1. Second, it is readily seen that for fixed r the action of the map ϑ_r is an expansion or contraction of the vector space $\mathbb{R}^3$ by the factor e^r. Thus, the ϑ_r invariance of the flow $\{\varphi_\Delta\}$ is the generalization of the invariance under scale change discussed in the Observations.

However, one might consider choosing a flow $\{\varphi_\Delta\}$ that models a motion from mechanics rather than one that commutes with ϑ_r . An example is the choice that φ_Δ is an isometry of $\mathbb{R}^3$. There are many reasons why such a restriction is not used, some of which will become apparent later. One of the reasons can be made precise immediately: it is that such a choice for φ_Δ can not in general be made in a manner that is consistent with the *sterance function* g . The following definition for sterance function is the obvious adaptation of the function g from Chapter 1.

4. DEFINITION. A *sterance* is a function $g\colon (\mathbb{R}^3 - \{0\}) \times I \longrightarrow \mathbb{R}$ that is smooth except on a closed subset of measure zero and that satisfies $\mathcal{R}(g) = 0$.

It is an immediate consequence that $g\circ\vartheta_r = g$, $r \in \mathbb{R}$, hence $\vartheta_r^* \, dg = dg$. (The ϑ_r invariance of g is the exact analog of the ψ_t invariance in Chapter 1.)

It follows from the definition that g is determined by $g|_{S^2\times I}$, the restriction of g to $S^2\times I$. For a fixed $t \in I$ we call $g|_{S^2\times\{t\}}$ the *image* associated with g at t ; we call $g|_{S^2\times I}$ the *image sequence*. In view of the fact that $g\circ\vartheta_r = g$, a distinguised role for $g|_{S^2\times I}$ is of significance only for applications: the image sequence $g|_{S^2\times I}$ is the function that is approximated in physical measurements.

5. DEFINITION. An admissible vector field X defined on a neighborhood $U \subset (\mathbb{R}^3-\{0\})\times I$ and a sterance g are said to be *compatible* if g is smooth on U and if $X(g) = 0$ on U .

An immediate consequence is that, on a domain of definition $J\times V$ of the local one parameter group $\{\varphi_\Delta\}$ of X , $\Delta \in (-\varepsilon,\varepsilon) = J \subset \mathbb{R}$, $V \subset U$, we have $g\circ\varphi_\Delta = g$, hence $\varphi_\Delta^* \, dg = dg$. This invariance of g along a path $\varphi_\Delta(m,t)$, $\Delta \in J$, is analogous to the Lambertian condition of Chapter 1.

It can now be seen that the condition that φ_Δ and ϑ_r commute is sufficient for the consistency of φ_Δ and ϑ_r with sterance. For, by the definition of sterance and by the definition of compatibility,

$$g\circ\vartheta_r\circ\varphi_\Delta = g\circ\varphi_\Delta = g = g\circ\vartheta_r = g\circ\varphi_\Delta\circ\vartheta_r \ .$$

If $dg - \frac{\partial g}{\partial s}ds \neq 0$ we have three linearly independent functions, g, $\ln\rho$, and s, on the 4-manifold $(\mathbb{R}^3-\{0\})\times I$. This is typical. The introduction of the sterance function always leaves us one function short of a chart for any of the manifolds used in vision problems. This gap is frequently noted and is often filled by changing g from a real valued function to a map into $\mathbb{R}^c$, $c \geq 2$. This is done typically to model the physical phenomena of color. (See for example Blicher 1985.) We shall not do this. Instead, we will make do with three functions, but from the three linearly independent 1-forms we will determine a fourth by the Hodge star operator. In this way we will have defined a bundle of bases on $\left\{dg - \frac{\partial g}{\partial s}ds \neq 0\right\}$.

A convenient definition of the Hodge star operator (Flanders 1963, 15-7) uses the metric induced on 1-forms by metric equivalence (O'Neill 1983, 60). Let M be an n-dimensional Riemannian manifold. Let $m \in M$. Metric equivalence refers to the existence of the isomorphism m from 1-forms at m to vectors at m defined by $\langle \mathrm{m}(\omega), X\rangle = \omega(X)$ for all 1-forms $\omega \in T^*_mM$ and all vectors $X \in T_mM$. The inner product of 1-forms ω_1 and ω_2 is defined by $\langle\omega_1, \omega_2\rangle = \langle \mathrm{m}(\omega_1), \mathrm{m}(\omega_2)\rangle$. For p-forms at m the inner product is defined by defining for decomposable elements $\nu = \nu_1\wedge\ldots\wedge\nu_p$ and $\mu = \mu_1\wedge\ldots\wedge\mu_p$, $\langle\nu, \mu\rangle = \det(\langle\nu_i, \mu_j\rangle)$. For M orientable, choose a volume element σ. (We assume a positive definite metric, hence $\langle\sigma,\sigma\rangle = 1$.) The Hodge star operator $*$ acting on a p-form ν produces an (n-p)-form $*\nu$ defined by

$$\nu\wedge\mu = \langle *\nu, \mu\rangle\, \sigma, \text{ for all (n-p)-forms } \mu .$$

From the smoothness of the metric tensor it follows that $*$ maps differential p-forms on M to differential (n-p)-forms on M .

We define a differential 1-form on $(\mathbb{R}^3-\{0\})\times I$ by

$$\beta = *\,(\,d\rho \wedge dg \wedge ds\,)\ .$$

We shall repeatedly use the following elementary facts about β . (The notation for the norm associated with the inner product of any of the vector spaces is $\|\ \| = \langle\ ,\ \rangle^{1/2}$.)

6. PROPOSITION.

1. $0 = \langle\beta, d\rho\rangle = \langle\beta, dg\rangle = \langle\beta, ds\rangle$.
2. $\|\beta\| = \|dg - \frac{\partial g}{\partial s}ds\|$.
3. $\vartheta_r^* \beta = \beta$.

Proof. Statement 1 is immediate. For example, $\langle\beta, d\rho\rangle\sigma = d\rho\wedge dg\wedge ds\wedge d\rho = 0$. For statement 2, if $dg - \frac{\partial g}{\partial s}ds = 0$ then we are done. Otherwise, by 1 we have that $d\rho$, $\dfrac{dg - \frac{\partial g}{\partial s}ds}{\|dg - \frac{\partial g}{\partial s}ds\|}$, ds , $\dfrac{\beta}{\|\beta\|}$ are orthonormal, hence $\sigma = d\rho \wedge \dfrac{dg - \frac{\partial g}{\partial s}ds}{\|dg - \frac{\partial g}{\partial s}ds\|} \wedge ds \wedge \dfrac{\beta}{\|\beta\|} = \dfrac{\langle\beta,\beta\rangle\ \sigma}{\|dg - \frac{\partial g}{\partial s}ds\|\ \|\beta\|}$.

To prove statement 3, we express β in spherical coordinates. For $x_1 = \rho\sin\theta\cos\phi$, $x_2 = \rho\sin\theta\sin\phi$, $x_3 = \rho\cos\theta$, $\sin\theta \neq 0$, with $dx_1\wedge dx_2\wedge dx_3\wedge ds = \sigma$, then $d\rho\wedge d\theta\wedge d\phi\wedge ds = \dfrac{\sigma}{\rho^2\sin\theta}$. Since $dg = \frac{\partial g}{\partial\theta}d\theta + \frac{\partial g}{\partial\phi}d\phi + \frac{\partial g}{\partial s}ds$, $\langle d\rho, d\rho\rangle = 1$, $\langle d\theta, d\theta\rangle = 1/\rho^2$, $\langle d\phi, d\phi\rangle = 1/(\rho^2\sin^2\theta)$, $\langle ds, ds\rangle = 1$, it follows that

$$\beta = \frac{1}{\sin\theta}\frac{\partial g}{\partial \phi}\, d\theta - \sin\theta \frac{\partial g}{\partial \theta}\, d\phi .$$

The desired result follows from $\theta\circ\vartheta_r = \theta$, $\phi\circ\vartheta_r = \phi$, $\frac{\partial g}{\partial \phi}\circ\vartheta_r = \frac{\partial g}{\partial \phi}$, and $\frac{\partial g}{\partial \theta}\circ\vartheta_r = \frac{\partial g}{\partial \theta}$. ■

It is convenient to have an indexed notation for these four 1-forms that are a basis for $T^*_{(m,t)}(\mathbb{R}^3-\{0\})\times I$ when $dg - \frac{\partial g}{\partial s}ds\Big|_{(m,t)} \neq 0$. We use

$$e^{1*} = dg \ , \quad e^{2*} = \beta \ , \quad e^{3*} = d\ln\rho \ , \quad e^{4*} = ds \ .$$

We denote the four vector fields that are dual to $\{e^{i*}\}$ by $\{e^i\}$. We also use $e_1 = e_g$, $e_2 = e_\beta$, $e_3 = e_\rho$, $e_4 = e_s$. It is easy to see that $e_g = \dfrac{\mathrm{grad} g - \frac{\partial g}{\partial s}\partial_s}{\|\mathrm{grad} g - \frac{\partial g}{\partial s}\partial_s\|^2}$, $e_\rho = \mathcal{R}$, and $e_s = \partial_s - \frac{\partial g}{\partial s}e_g$ by checking that $e^{i*}(e_j) = \delta_{ij}$.

7. PROPOSITION.

1. $\|e_g\| = \dfrac{1}{\|\beta\|} = \|e_\beta\|$.
2. $d\vartheta_r\, e_i = e_i$, hence $[\mathcal{R}, e_i] = 0$, $i=1,2,3,4$.

Proof. The first equation in statement 1 follows fron the formulas for e_g and statement 2 in Proposition 6. For the second equation, note that statement 1 of Proposition 6 implies $m(\beta)$ = constant$\cdot e_\beta$.

Statement 2 follows from the fact that $\vartheta_r^*\, e^{i*} = e^{i*}$, $i=1,2,3,4$.

■

Note that three of the 1-forms are differentials of functions. However $d\beta$ need not vanish. From this observation we have

8. PROPOSITION. $[e_i, e_j] = \beta([e_i, e_j])\, e_\beta\,, \quad i, j = 1,2,3,4\,.$

The following table summarizes the definitions, notation, and results of this section.

TABLE 1

functions		g		ρ $\rho(m,t)=\lvert m\rvert$	s
vector fields	X			$\mathcal{R}=\sum_i x_i \frac{\partial}{\partial x_i}$	$\partial_s = \text{grad}\, s$
flows	φ_Δ (local)			ϑ_r	
1-forms ↑ dual ↓		$e^{1*}=dg$	$e^{2*}=\beta$	$e^{3*}=d\ln\rho$	$e^{4*}=ds$
vector fields		$e_1=e_g = \dfrac{\text{grad}\, g - \frac{\partial g}{\partial s}\partial_s}{\lVert \text{grad}\, g - \frac{\partial g}{\partial s}\partial_s \rVert}$	$e_2=e_\beta$	$e_3=\mathcal{R}$	$e_4=\partial_s - \frac{\partial g}{\partial s} e_g$
relations	$\varphi_\Delta \circ \vartheta_r = \vartheta_r \circ \varphi_\Delta$ $[X, \mathcal{R}]=0$	$g \circ \vartheta_r = g$ $g \circ \varphi_\Delta = g$		$\vartheta_r^*\, e^{i*} = e^{i*}$ $d\vartheta_r\, e_i = e_i$ $[\mathcal{R}, e_i]=0$	

$$\gamma_{ij}=\langle e_i,e_j\rangle= \begin{bmatrix} \|e_g\|^2 & 0 & 0 & -\frac{\partial g}{\partial s}\|e_g\|^2 \\ 0 & \|e_\beta\|^2 & 0 & 0 \\ 0 & 0 & \rho^2 & 0 \\ -\frac{\partial g}{\partial s}\|e_g\|^2 & 0 & 0 & 1+\left(\frac{\partial g}{\partial s}\right)^2\|e_g\|^2 \end{bmatrix}$$

It is easy to see from Table 1 that for any sterance g there exists admissible vector fields that are compatible with g. The simplest example is the vector field e_s. Another is $\mathcal{R} + e_s$. More precisely, $X = \beta(X)\, e_\beta + d\ln\rho(X)\, \mathcal{R} + e_s$ is admissible if and only if $\mathcal{R}\big(\beta(X)\big) = 0$ and $\mathcal{R}\big(d\ln\rho(X)\big) = 0$, for then $[X,\mathcal{R}] = 0$. If this X is admissible, then it is clearly compatible with g. This is summarized in the following.

9. PROPOSITION. Let g be a sterance that is smooth on $U \subset (\mathbb{R}^3-\{0\})\times I$ and assume $dg - \frac{\partial g}{\partial s}ds \neq 0$ on U. Then there exists vector fields on U that are compatible with g.

To say that X is compatible with g is to say that the motion $\varphi_\vartriangle(m,t)$, $\vartriangle \in J$, of a point (m,t) is contained in a level set of g (Figure 7). Equivalently, g is constant along $\varphi_\vartriangle(m,t)$. This latter is the usual point of view for applications: the sterance associated with the point (m,t) does not change as the point moves. In this latter sense g is a consequence of a motion. This is analagous to our consideration in Chapter 1 of a form $gi_g{}^*\sigma$ that is a consequence of a

Lambertian form $f\nu$ (Propositions 27 and 28). Just as in Chapter 1, to get a well posed problem we do not place such restrictions on g. As in Chapter 1, g has essentially no restrictions. Instead, we will study subsets of admissible vector fields.

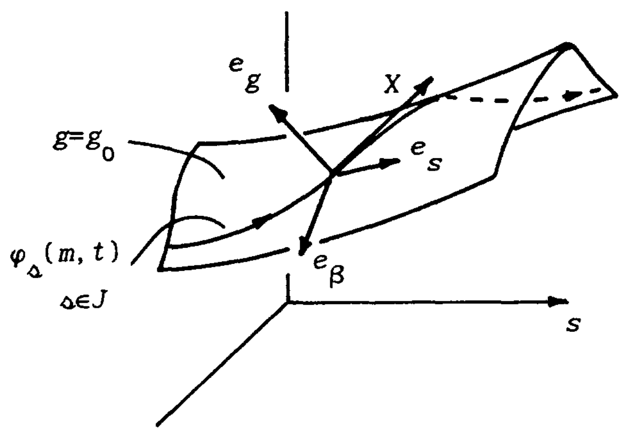


Figure 7

In this section there has been no mention of objects or surfaces. In fact, there will be no mention of these until Section 4. The reason for this was suggested in the first section but can now be made precise. Let S be a neighborhood in any surface such that $\mathcal{R}$ is nowhere tangent to S. Then $\{\vartheta_r(S)\}_r$ is a family of surfaces and for each surface its motion due to admissible, compatible X is pointwise along

level sets of g.

In applications this situation is frequently stated in the following manner: the absolute distance from the origin to a moving surface cannot be determined from $g|_{S^2 \times I}$, where the sterance g is the consequence of a Lambertian form on the surface.

In light of this it is not meaningful to introduce a surface. However, $\{\vartheta_r(S)\}_r$ suggests that we introduce $\{d\vartheta_r(TS)\}_r$, that is, an involutive distribution, and that we seek relationships between X , g , and integral manifolds of this distribution.

2.3 Parallel fields, geodesic variations in images

In The last section we saw that admissible, compatible vector fields always exist but are not uniquely determined by the sterance. In this section we will consider a subset of admissible vector fields. In this restricted subset we obtain a uniqueness result, up to images which satisfy a condition related to geodesic variations on S^2. When this uniqueness result is applied to the case of a constant image sequence, i.e., $\frac{\partial g}{\partial s} = 0$, we find that this restricted problem satisfies the still picture condition.

The following terminology is standard (for example, O'Neill 1983, Ch.3). For vector fields V and W on $(\mathbb{R}^3-\{0\})\times I$, let $D_V W$ denote the natural covariant derivative of W with respect to V ,

$$D_V W = \sum_i V(dx_i W) \frac{\partial}{\partial x_i} + V(ds\ W) \frac{\partial}{\partial s} ,$$

where x_1, x_2, x_3, s are the natural coordinate functions. A vector field P is said to be parallel if $D_Y P = 0$ for all vector fields Y . A related case is $ds\ P = 0$ and, for all Y such that $ds\ Y = 0$, $D_Y P = 0$. That is, P is tangent to and parallel in $\mathbb{R}^3-\{0\}\times\{t\}$, but $D_{\partial_s} P$ need not vanish.

Let X be an admissible vector field. Let X and sterance g be compatible on a neighborhood U (where we may assume $U = \bigcup_{r\in\mathbb{R}} \vartheta_r(U)$) in which $dg - \frac{\partial g}{\partial s}ds \neq 0$. In this case we have on U the four linearly independent vector fields $e_1=e_g$, $e_2=e_\beta$, $e_3=\mathcal{R}$, $e_4=e_s$, and the dual

1-forms dg , β , $d\ln\rho$, ds . Therefore $X = \beta(X)\, e_\beta + d\ln\rho(X)\, \mathcal{R} + e_s$, since $dg\, X = 0$ and $ds\, X = 1$.

10. THEOREM. Let X and X' be admissible vector fields on $U \subset (\mathbb{R}^3-\{0\})\times I$, and let g be a sterance such that both X and X' are compatible with g on U. If there exists a vector field $P \neq 0$ on U , with $D_Y P = 0$ for all Y with $ds\, Y = 0$, and a smooth function h on U such that $X' - X = hP$, where the measure of $\{h=0\}$ is zero, then

$$\beta(P)\ \beta\big(D_{e_\beta} e_g\big) = 0 \qquad (1)$$

and

$$\beta(P)\ \beta\big(D_{e_g} e_g\big) = d\ln\rho(P) \quad . \qquad (2)$$

Proof. It follows from the admissibilty and compatibility of X and X' that $hP = d\ln\rho(hP)\, \mathcal{R} + \beta(hP)\, e_\beta$. From the matrix for $<e_i, e_j>$ in Table 1 of Section 2, it follows that $<hP, e_g> = <hP, e_s> = 0$ on U . Let $e_1=e_g$ and $e_4=e_s$. It follows by continuity and from the fact that $\{h=0\}$ has measure zero that

$$<P, e_j> = 0 \text{ on } U\ ,\ j=1,4\ . \qquad (a)$$

Since $Y<P, e_j> = <D_Y P, e_j> + <P, D_Y e_j>$, we have for every vector field Y with $ds\, Y = 0$,

$$<P, D_Y e_j> = 0 \text{ on } U\ ,\ j=1,4\ . \qquad (b)$$

Note that $D_Y \mathcal{R} = Y$ for any vector field Y . Recall $e_3 = \mathcal{R} = \rho\, \mathrm{grad}\rho$. Hence,

$$\rho d\rho(D_Y e_j) = <\mathcal{R}, D_Y e_j> = Y<e_3, e_j> - <D_Y \mathcal{R}, e_j> = -<Y, e_j>\ ,\ j=1,2,3,4\ . \qquad (c)$$

Since $D_Y e_j = d\ln\rho(D_Y e_j)\ \mathcal{R} + \beta(D_Y e_j)\ e_\beta$ + (terms in e_s , e_g) , it follows from (a), (b), and (c) that

$$0 = \langle P, D_Y e_j\rangle = -d\ln\rho(P)\ \langle Y, e_j\rangle \frac{\|\mathcal{R}\|^2}{\rho^2} + \beta(P)\ \beta\big(D_Y e_j\big)\ \|e_\beta\|^2\ ,\ j=1,4\ . \quad (*)$$

Recall $\|\mathcal{R}\|^2 = \rho^2$ and $\|e_\beta\|^2 = \|e_g\|^2$. Note $\|e_g\| \neq 0$. (See Table 1.) From (*) with j=1 and with $Y=e_g$ it follows that

$$0 = -d\ln\rho(P) + \beta(P)\ \beta\big(D_{e_g} e_g\big)\ ;$$

with $Y=e_\beta$ it follows that

$$0 = \quad 0 \quad + \beta(P)\ \beta\big(D_{e_\beta} e_g\big)\ .$$

(Note that (*) is trivial for $Y=\mathcal{R}$. Also (*) with j=4 leads to the same results, for $e_4 = e_s = \partial_s - \frac{\partial g}{\partial s} e_g$ and ∂_s is parallel.) ∎

It is easy to see that if $hP = X' - X$, then $\mathcal{R}(h) = h$, for $[\mathcal{R}, hP]$ must vanish if X and X' are admissible, and $[\mathcal{R}, P] = D_{\mathcal{R}}P - D_P\mathcal{R} = -P$. Consequently, there exists admissible vector fields $X + hP$. An admissible vector field $\partial_s + hP$ is illustrated in Figure 8.

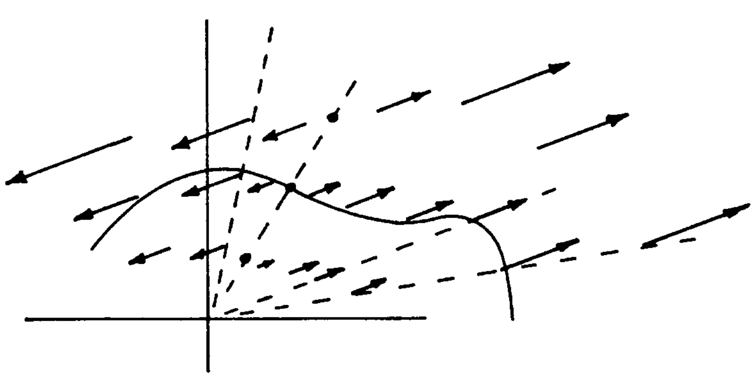


Figure 8

The following corollaries explain why the theorem is interesting. Note that $U \cap s^{-1}(t) = U \cap \mathbb{R}^3-\{0\}\times\{t\}$. The corollaries are direct consequences of statements (1) and (2) of the theorem.

11. COROLLARY. For P as in the theorem, if there exists a point $(m,t) \in U \cap s^{-1}$ such that $\beta(P)_{(m,t)} = 0$, then $P = 0$ on $U \cap s^{-1}$.

Proof. Apply statement (2) of the theorem and $0 = dg\ P = ds\ P$.■

12. EXAMPLE. Let $\frac{\partial g}{\partial s} = 0$ on U . Hence, on U the image sequence $g|_{S^2\times I}$ is a still picture, $g|_{S^2\times I} = g|_{S^2\times\{t\}}$. Let $X = e_S$, hence $X = \partial_s$. Thus $\varphi_\Delta = id \times \{t \longmapsto t+\Delta\}$. The corollary tells us that $X + hP = \partial_s + hP$, $P \neq 0$, cannot be compatible with a still picture if any point $(m,t) \in U$ has $\beta(P)_{(m,t)} = 0$. Note that it is not necessary that $h(m,t) \neq 0$.

The vector field $\partial_s + hP$ with $h \neq 0$ is frequently implicit in the engineering literature. Compare the problem of finding the "focus of expansion," that is, a point (m,t) such that $\beta(m,t) = 0$, in Prazdny (1983).

The proofs of the next two corollaries are immediate.

13. COROLLARY. With P as in the theorem, if there exists a point $(m,t) \in U \cap s^{-1}(t)$ such that $d\rho\ P = 0$ and $\beta\left(D_{e_g} e_g\right) \neq 0$, then $P = 0$ on $U \cap s^{-1}(t)$.

The following result is of more significance than the preceding two for it depends on only the image $g|_{S^2\times\{t\}}$.

14. COROLLARY. With P as in the theorem, if there exists a point $(m,t) \in U \cap s^{-1}(t)$ such that $\beta\left(D_{e_\beta} e_g\right)_{(m,t)} \neq 0$, then $P = 0$ on $U \cap s^{-1}(t)$.

This result is of the type we are seeking. First, it is local: $\beta\left(D_{e_\beta} e_g\right)$ depends on only the germ of g at (m,t) . Second, it provides a uniqueness result. For example, if $\frac{\partial g}{\partial s} = 0$ on U and $\beta\left(D_{e_\beta} e_g\right)_{(m,t)} \neq 0$ for any $(m,t) \in U$, then there is a neighborhood $\bigcup_{|t|<\varepsilon} U \cap s^{-1}(t)$ on which $\partial_s + hP$ is a solution if and only if $P \equiv 0$.

The following corollary provides some insight into what the image $g|_{S^2\times\{t\}}$ must look like locally if $P \neq 0$. Recall the definition of geodesic variation. Here we follow O'Neill (1983, Ch.8). A two parameter map x into S^2 is a map $x: \mathcal{D} \longrightarrow S^2$, where $\mathcal{D}$ is open in $\mathbb{R}^2$ and where horizontal and vertical lines in $\mathbb{R}^2$ intersect $\mathcal{D}$ in intervals. Let (u,v) be the natural coordinates for $\mathbb{R}^2$. The u-parameter curves $u \longmapsto x(u,v_0)$ are called longitudinal. A two parameter map x is a geodesic variation if every longitudinal curve of x is a geodesic. With ∂_u , ∂_v the natural coordinate tangent vectors of $\mathcal{D}$, there are the two vector fields along x , $dx(\partial_u)$ and $dx(\partial_v)$.

Recall that the covariant derivative D^{T} on $S^2\times\{t\}$ induced from D on $(\mathbb{R}^3-\{0\})\times I$ is defined by taking the tangential component, and a

curve γ is a geodesic on $S^2\times\{t\}$ if and only if $D^{\mathsf{T}}_{\dot\gamma}\ \dot\gamma = 0$. Recall that e_g and e_β are tangent to $S^2\times\{t\}$.

15. COROLLARY. For P as in the theorem, $P \neq 0$ if and only if for every $(m,t) \in \{P\neq 0\}\cap U$ there exists a geodesic variation $x\colon [-\varepsilon,\varepsilon]\times(-\delta,\delta) \longrightarrow S^2 \hookrightarrow (\mathbb{R}^3-\{0\})\times\{t\}$ such that $x(0,0) = (m,t)$ and

$$dx(\partial_u) = \frac{e_\beta}{\|e_\beta\|} \quad , \qquad dx(\partial_v) = e_g \quad .$$

Consequently, $P \neq 0$ if and only if

$$D^{\mathsf{T}}_{\left(\frac{e_\beta}{\|e_\beta\|}\right)}\frac{e_\beta}{\|e_\beta\|} = 0 \quad \text{and} \quad D^{\mathsf{T}}_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} D^{\mathsf{T}}_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} e_g = \mathbf{R}_{e_g\left(\frac{e_\beta}{\|e_\beta\|}\right)}\left(\frac{e_\beta}{\|e_\beta\|}\right) ,$$

that is, e_g is a Jacobi vector field, where $\mathbf{R}$ is the curvature tensor for S^2.

Proof. For P as in the theorem, $P \neq 0$ on $U \cap s^{-1}(t)$ if and only if $\beta\big(D_{e_\beta} e_g\big) = 0$ on $U \cap s^{-1}(t)$. Since $D_V W$ is tensorial in V,

$$\beta\Big(D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} e_g\Big) = 0 \quad . \tag{$*$}$$

Also, $< \frac{e_\beta}{\|e_\beta\|} , \frac{e_\beta}{\|e_\beta\|} > = 1$ implies $\beta\Big(D_{e_g}\frac{e_\beta}{\|e_\beta\|}\Big) = 0$. Hence

$$0 = \beta\left[D_{e_g}\frac{e_\beta}{\|e_\beta\|} - D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} e_g\right] = \beta\left(\left[\ e_g\ ,\ \frac{e_\beta}{\|e_\beta\|}\ \right]\right) .$$

Since $[e_g\ ,\ e_\beta] = \beta\big([e_g\ ,\ e_\beta]\big)e_\beta$, we have $\left[e_g\ ,\ \frac{e_\beta}{\|e_\beta\|}\right] = 0$.

Therefore, considering e_g and $\frac{e_\beta}{\|e_\beta\|}$ as tangent vectors of $S^2\times\{t\}$, for

each point of $S^2\times\{t\} \cap U$ there exists a coordinate neighborhood about the point such that e_g and $\frac{e_\beta}{\|e_\beta\|}$ are coordinate vector fields.

From $< \frac{e_\beta}{\|e_\beta\|} , \frac{e_\beta}{\|e_\beta\|} > = 1$ we have $< \frac{e_\beta}{\|e_\beta\|} , D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} \frac{e_\beta}{\|e_\beta\|} > = 0$.

From $< \frac{e_\beta}{\|e_\beta\|} , e_g > = 0$ we have

$$< e_g , D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} \frac{e_\beta}{\|e_\beta\|} > + < \frac{e_\beta}{\|e_\beta\|} , D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} e_g > = 0 \quad .$$

From (*) and Table 1, $< \frac{e_\beta}{\|e_\beta\|} , D_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} e_g > = 0$. Consequently, $D^{\mathrm{T}}_{\left(\frac{e_\beta}{\|e_\beta\|}\right)} \frac{e_\beta}{\|e_\beta\|} = 0$, so that the coordinate corresponding to the coordinate field $\frac{e_\beta}{\|e_\beta\|}$ is in fact a geodesic curve on $S^2\times\{t\}$. Thus the coordinates define a geodesic variation and the transverse field e_g satisfies the Jacobi differential equation (e.g., O'Neill 1983, Lemma 8.3). ∎

The results of this section are, of course, only partial. However, they do begin to clarify several relationships between admissible, compatible fields and the image. We may summarize these results by saying that for a sterance g , for $X + hP$ admissible and compatible with g , and for P parallel on $(\mathbb{R}^3-\{0\})\times\{t\}$ for all t , with $ds\ P = 0$, we understand the necessary and sufficient conditions on g such that $\left\{P \mid \left(X+hP\right)(g)=0\right\} \neq \left\{P=0\right\}$.

2.4 Isometries for two-dimensional distributions

The results of the last section described a restricted subset of admissible vector fields and used only the properties of the sterance function. In this section a larger class of admissible vector fields is studied by the introduction of additional structure. As suggested at the end of Section 2, the new structure is involutive distributions or, equivalently, differential ideals (Warner 1971, Ch. 1, 2; Bishop and Goldberg 1968, §3.11). The vector fields to be studied are those whose flows are isometries of the integral manifolds of the distributions. Our first task is to define the differential ideal that is of interest.

16. DEFINITION. A 1-form compatible with an admissible vector field X is a closed 1-form ω defined on a neighborhood $U \subset (\mathbb{R}^3-\{0\})\times I$ such that

i. about each point $(m,t) \in U$ there is a neighborhood $V = \bigcup_{r\in\mathbb{R}} \vartheta_r(V)$ in which $\omega = dF$, $F \subset C^\infty(V)$;

ii. $\omega(\mathcal{R}) = 1$;

iii. $\omega(X) = 0$.

This definition merely summarizes the situation in which there is a two-dimensional submanifold (surface) in $(\mathbb{R}^3-\{0\})\times\{t_0\}$, with $\mathcal{R}$ never tangent. Consequently, there is a coordinate chart $\mu = (y_1,y_2,y_3,s)$ on

a neighborhood such that $(0, y_2, y_3, t_0)$ is a coordinate chart for the surface. And, since $[X, \mathcal{R}] = 0$, the inverse of the map $(r, y_2, y_3, \vartriangle) \longmapsto \vartheta_r \circ \varphi_\vartriangle \circ \mu^{-1}(0, y_2, y_3, t_0)$ is a coordinate chart on a neighborhood. Let F be the first coordinate function of this inverse. Each level set of F is a submanifold which is tangent to X .

On the other hand, directly from the definition the 1-form ω trivially generates a differential ideal. Hence, by the Frobenius theorem (Warner 1971, Theorem 2.32), through each point (m, t) of U there is a unique, maximal, connected codimension one submanifold whose tangent space is the annihilator of ω .

There are the following immediate consequences of the definition: $\vartheta_r^* \omega = \omega$, hence $L_\mathcal{R} \omega = 0$ where $L_\mathcal{R}$ is the Lie derivative relative to $\mathcal{R}$; $F \circ \varphi_\vartriangle = F$, hence $\varphi_\vartriangle^* \omega = \omega$ and $L_X \omega = 0$. Further, recall, as in Table 1, that $\vartheta_r^* \, e^{1*} = e^{1*}$, $[\mathcal{R}, X] = 0$, hence

17. PROPOSITION. In the expansion $\omega = \sum_i f_i e^{i*}$, where ω is a 1-form that is compatible with an admissible vector field X,

$$f_3 = 1 \quad \text{and} \quad \mathcal{R}(f_i) = 0 \ , \ i=1,2,3,4 \ .$$

To construct a basis we use the Hodge star operator.

18. PROPOSITION. For a sterance g, $dg - \frac{\partial g}{\partial s} ds \neq 0$ on $U = \bigcup_{r \in \mathbb{R}} \vartheta_r(U)$, for a 1-form ω compatible with an admissible vector field X defined on U, then the differential 1-forms

$$a^{1*} = dg\ ,$$
$$a^{2*} = \nu = *\ (\rho\omega \wedge dg \wedge ds)\ ,$$
$$a^{3*} = \omega\ ,$$
$$a^{4*} = ds\ ,$$

constitute a linearly independent set of 1-forms on U , with

$$\nu = f_3\beta - \rho^2 f_2 \|\beta\|^2 d\ln\rho\ , \quad f_2 = \omega(e_2)\ , \quad f_3 = \omega(e_3) = 1\ ,$$

and

$$\vartheta_r^* \nu = \nu\ ,$$

hence,

$$\vartheta_r^* a^{i*} = a^{i*}\ ,\ i=1,2,3,4\ , \quad \text{but} \quad \varphi_\Delta^* a^{i*} = a^{i*}\ ,\ i\neq 2\ .$$

Proof. It follows from $\omega = \sum_i f_i\ e^{i*}$ and from the definition of β (see Table 1) that $\nu = *\ (\rho f_2\beta \wedge dg \wedge ds) + f_3\beta$. By the definition of the Hodge star operator

$$\langle\ (\nu - f_3\beta)\ ,\ e^{i*}\ \rangle\ \sigma \ = \ \rho f_2(\beta \wedge dg \wedge ds \wedge e^{i*}) = \begin{cases} 0 & \text{if } i\neq 3 \\ -\ f_2\ d\rho\wedge dg\wedge ds\wedge\beta = -\ f_2\ \|\beta\|^2\ \sigma \end{cases}$$

The result follows from $\|e^{3*}\|^2 = \|d\ln\rho\|^2 = 1/\rho^2$.

It is a corollary to the proof of Proposition 6 that $\rho\|\beta\|$ is independent of ρ, hence $\vartheta_r^*\ \nu = \nu$. The remaining relations are previous results and definitions. ∎

19. PROPOSTION. The properties of $\{a^{i*}\}$ and the dual vectors $\{a_i\}$ for a neighborhood $U = \bigcup_{r\in\mathbb{R}} \vartheta_r(U) \subset (\mathbb{R}^3-\{0\})\times I$ are collected in the

following table, where

$$D = \det(a^{i*}(e_j)) = f_3{}^2 + f_2{}^2\rho^2\|\beta\|^2 \ ,$$

$$f_3 = 1 \quad , \quad \text{and} \quad \|e_g\| = \|e_\beta\| = 1/\|\beta\| \ .$$

TABLE 2

$$a^{1*} = dg \qquad a_1 = a_g = e_g - f_1 a_\omega$$

$$a^{2*} = \nu \qquad a_2 = a_\nu = \frac{1}{D}\left(f_3 e_\beta - f_2 \mathcal{R}\right) = \frac{1}{f_3}\left(e_\beta - f_2 a_\omega\right)$$

$$a^{3*} = \omega = \sum_i f_i e^{i*} \qquad a_3 = a_\omega = \frac{1}{D}\left(\rho^2 f_2{}^2\|\beta\|^2 e_\beta + f_3 \mathcal{R}\right)$$

$$a^{4*} = ds \qquad a_4 = a_s = e_s - f_4 a_\omega$$

$$\vartheta_r^* \, a^{i*} = a^{i*} \ , \quad L_{\mathcal{R}} a^{i*} = 0 \ , \qquad d\vartheta_r \, a_i = a_i \ , \quad [\mathcal{R}, a_i] = 0 \ , \ i=1,2,3,4$$

$$\varphi_\Delta^* \, a^{i*} = a^{i*} \ , \quad L_X a^{i*} = 0 \ , \qquad d\varphi_\Delta \, a_i = a_i + \left(\varphi_\Delta^* \, \nu\right)(a_i) \ a_\nu \ , \ i \neq 2$$

$$d\varphi_\Delta \, a_\nu = \left(\varphi_\Delta \, \nu\right)(a_\nu) \ a_\nu$$

$$[X, a_i] = -\left(L_X \nu\right)(a_i) \ a_\nu \ , \ i=1,2,3,4$$

$$\langle a_2 \, , \, a_i \rangle = 0 \ , \quad i \neq 2$$

In particular, at each point in U the vectors $a_1 = a_g$ and $a_2 = a_\nu$ form an orthogonal basis for the integral manifold $\left\{\omega=0 \ , \ ds=0\right\}$, and under the flow $\{\varphi_\Delta\}$ of X the basis a_g , a_ν is mapped to a basis.

Proof. The expressions for a_i in terms of the basis $\{e_i\}$ is obtained by inverting the matrix $\left(a^{i*}(e_j)\right)$. Note that $\vartheta_r^* \; a^{i*} = a^{i*}$ implies $L_{\mathcal{R}}a_i^* = 0$ (Warner 1971, Prop. 2.25) and that the converse holds locally. The same holds for $d\vartheta_r a_i = a_i$ and $[\mathcal{R}, a_i] = L_{\mathcal{R}}a_i$. The results for $d\varphi_{\vartriangle} a_i$ follow from $a^{j*}\left(d\varphi_{\vartriangle} a_i\right) = \varphi_{\vartriangle}^* \; a^{j*}(a_i)$. Similarly for $d\vartheta_r a_i$. The result for $[X, a_i]$ follows most simply from the product rule for $X\left(a^{j*}(a_i)\right)$ and from $a^{j*}\left([X, a_i]\right) = 0$, $j \neq 2$, since only ν is not exact and since $a^{j*}(X)$ is a constant, $j \neq 2$. (It also follows from the derivative of $d\varphi_{\vartriangle} a_i$ with respect to $\vartriangle$.) The orthogonality result follows directly by using $<e_i \; , \; e_j>$ in Table 1. ∎

We now reach the motivation for the preceding constructions. We wish to study X , g , and ω as in the following definition.

20. DEFINITION. Let X be an admissible vector field with local one parameter group $\varphi_{\vartriangle}$, where $\varphi_{\vartriangle}$ is defined on U for $\vartriangle \in J$ $(0 \in J)$. Let g be a sterance, $dg - \frac{\partial g}{\partial s}ds \neq 0$, which is compatible with X , and let ω be a 1-form which is compatible with X . We say $\varphi_{\vartriangle}$ is an isometry of $\left\{\omega=0 \; , \; ds=0\right\}$ if $\varphi_{\vartriangle}$ is an isometry from the integral manifolds of $\left\{\omega=0 \; , \; ds=0\right\} \cap U$ to the integral manifolds of $\left\{\omega=0 \; , \; ds=0\right\} \cap U$. That is, $\varphi_{\vartriangle}$ is an isometry if for each $(m, t) \in U$ and $\vartriangle \in J$

$$< d\varphi_{\vartriangle} a_i \; , \; d\varphi_{\vartriangle} a_j >_{\varphi_{\vartriangle}(m,t)} = < a_i \; , \; a_j >_{(m,t)} \; , \quad i,j = 1,2,$$

where $a_1 = a_g$ and $a_2 = a_\nu$ as defined in Proposition 19.

The following result provides several equivalent descriptions of an isometry of $\left\{\omega=0\ ,\ ds=0\right\}$.

21. THEOREM. For a local one parameter group φ_Δ , defined on U for $\Delta \in J$, of an admissible vector field X , X compatible with sterance g , $dg - \frac{\partial g}{\partial s}ds \neq 0$, and compatible with 1-form ω , the following are equivalent. For $\Delta \in J$, $(m,t) \in U$:

1. φ_Δ is an isometry of $\left\{\omega=0\ ,\ ds=0\right\}$.

2. $$X\langle a_g\ ,\ a_g\rangle = 0$$
$$d\varphi_\Delta\, a_g = a_g$$
$$d\varphi_\Delta\left(\frac{a_\nu}{\|a_\nu\|}\right) = \frac{a_\nu}{\|a_\nu\|}$$

3. $$X\langle a_g\ ,\ a_g\rangle = 0$$
$$\left[X\ ,\ a_g\right] = 0$$
$$\left[X\ ,\ \frac{a_\nu}{\|a_\nu\|}\right] = 0$$

4. $$\langle D_{a_g}X\ ,\ a_g\rangle = 0$$
$$\langle D_{a_g}X\ ,\ \frac{a_\nu}{\|a_\nu\|}\rangle + \langle D_{\left(\frac{a_\nu}{\|a_\nu\|}\right)}X\ ,\ a_g\rangle = 0$$
$$\langle D_{\left(\frac{a_\nu}{\|a_\nu\|}\right)}X\ ,\ \left(\frac{a_\nu}{\|a_\nu\|}\right)\rangle = 0$$

5. For every $V \in \left\{\omega=0\ ,\ ds=0\right\}$, $\langle D_V X\ ,\ V\rangle = 0$

6. The (0,2) tensor $T_X \in T_2^0\left(\left\{\omega=0\ ,\ ds=0\right\}\right)$ defined by

$$T_X(V,W) = \langle D_V X\ ,\ W\rangle$$

is skew symmetric.

Proof. 1 $\Longleftrightarrow$ 2. From Proposition 19

$$d\varphi_\Delta\, a_g = a_g + \left(\varphi_\Delta^*\, \nu\right)(a_g)\; a_\nu \qquad (*)$$

$$d\varphi_\Delta\, a_\nu = \left(\varphi_\Delta^*\, \nu\right)(a_\nu)\; a_\nu \qquad (**)$$

($\Rightarrow$) If φ_Δ is an isometry, then $(\varphi_\Delta^*\, \nu)(a_g) = 0$ since a_g and a_ν form an orthogonal basis. If φ_Δ is an isometry $\langle d\varphi_\Delta a_g\, ,\, d\varphi_\Delta a_g\rangle_{\varphi_\Delta(m,t)} = \langle a_g\, ,\, a_g\rangle_{(m,t)}$. But $(\varphi_\Delta^*\nu)(a_g)=0$ implies $\langle d\varphi_\Delta a_g\, ,\, d\varphi_\Delta a_g\rangle_{\varphi_\Delta(m,t)} = \langle a_g\, ,\, a_g\rangle_{\varphi_\Delta(m,t)}$ by $(*)$. Hence $\langle a_g\, ,\, a_g\rangle_{\varphi_\Delta(m,t)}$ is constant for $\Delta \in J$, thus $X\langle a_g\, ,\, a_g\rangle = 0$. Finally, if φ_Δ is an isometry, then $1 = \left\langle d\varphi_\Delta\left[\frac{a_\nu}{\|a_\nu\|}\right]\, ,\, d\varphi_\Delta\left[\frac{a_\nu}{\|a_\nu\|}\right]\right\rangle_{\varphi_\Delta(m,t)}$. But by $(**)$ $d\varphi_\Delta\left[\frac{a_\nu}{\|a_\nu\|}\right]_{(m,t)} = (\varphi_\Delta^*\nu)_{(m,t)}\left[\frac{a_\nu}{\|a_\nu\|}\right]_{(m,t)}(a_\nu)_{\varphi_\Delta(m,t)}$. Hence $(\varphi_\Delta^*\nu)_{(m,t)}\left[\frac{a_\nu}{\|a_\nu\|}\right]_{(m,t)} = \pm\frac{1}{\|a_\nu\|_{\varphi_\Delta(m,t)}}$. By continuity at $\Delta = 0$ the right hand side must be positive. Hence by $(**)$ $d\varphi_\Delta\left[\frac{a_\nu}{\|a_\nu\|}\right]_{(m,t)} = \left[\frac{a_\nu}{\|a_\nu\|}\right]_{\varphi_\Delta(m,t)}$.

($\Leftarrow$) $\langle d\varphi_\Delta a_g\, ,\, d\varphi_\Delta a_g\rangle_{\varphi_\Delta(m,t)} = \langle a_g\, ,\, a_g\rangle_{(m,t)}$ follows from $X\langle a_g\, ,\, a_g\rangle = 0$ and from $d\varphi_\Delta a_g = a_g$. That $d\varphi_\Delta$ preserves orthogonality is easy to see. Finally, $\langle d\varphi_\Delta a_\nu\, ,\, d\varphi_\Delta a_\nu\rangle_{\varphi_\Delta(m,t)} = \|a_\nu\|^2_{(m,t)}$ follows directly.

2 $\Longleftrightarrow$ 3. The forward ($\Rightarrow$) case follows directly from the definition of Lie derivative that uses the derivative of $d\varphi_\Delta a_1$ with respect to Δ . Conversely, the vanishing of the Lie bracket implies that locally the

corresponding local one parameter groups commute, hence the result.

$3 \Leftrightarrow 4$. This follows from the identity $D_X Y - D_Y X = [X,Y]$ and from $a^{i*}([X,a_j]) = 0$, $i \neq 2$.

$$<D_{a_g} X \, , \, a_g> = <D_X a_g + [X,a_g], a_g> = (1/2)X<a_g \, , a_g> \, ;$$

$$<D_{a_g} X \, , \, \frac{a_\nu}{\|a_\nu\|}> + <D_{\left(\frac{a_\nu}{\|a_\nu\|}\right)} X \, , \, a_g>$$

$$= <D_X a_g + [X,a_g] \, , \, \frac{a_\nu}{\|a_\nu\|}> + <D_X\left(\frac{a_\nu}{\|a_\nu\|}\right) + \left[X, \frac{a_\nu}{\|a_\nu\|}\right] \, , \, a_g>$$

$$= X<a_g \, , \, \frac{a_\nu}{\|a_\nu\|}> + <[X,a_g] \, , \, \frac{a_\nu}{\|a_\nu\|}> \;=\; <[X,a_g] \, , \, \frac{a_\nu}{\|a_\nu\|}> \, ;$$

$$<D_{\left(\frac{a_\nu}{\|a_\nu\|}\right)} X \, , \, \frac{a_\nu}{\|a_\nu\|}> = <\left[X \, , \, \frac{a_\nu}{\|a_\nu\|}\right] \, , \, \frac{a_\nu}{\|a_\nu\|}> \, .$$

$4 \Leftrightarrow 5$. 4 clearly is a special case of 5. And 5 follows from 4 by the tensorial property of the covariant derivative and the fact that a_g and a_ν constitute a basis for the tangent space of the integral manifolds.

$5 \Leftrightarrow 6$. This requires only the application of the polarization identity for bilinear forms. ■

A key reason for considering only those admissible vector fields which are isometries of $\left\{\omega=0 \, , \, ds=0\right\}$ is that this restriction excludes the type of motion in which the still picture condition is violated by the retraction of $\mathbb{R}^3$ to the origin. (Figure 9)

22. COROLLARY. (radial motion) Let X be an admissible vector field on $U \subset (\mathbb{R}^3-\{0\})\times I$ with local one parameter group φ_Δ , and let X be compatible with a sterance g , $dg - \frac{\partial g}{\partial s}ds \neq 0$, and with 1-form ω . Let φ_Δ be an isometry of $\left\{\omega=0\ ,\ ds=0\right\}$. If $X = X_3\mathcal{R} + \partial_s$ on U , $X_3 \in C^\infty(U)$, then $X_3 = 0$ on U .

Proof. Let V be any vector in $\left\{\omega=0\ ,\ ds=0\right\}$ at $(m,t) \in U$ such that $dX_3(V) = 0$. Then, by statement 5 of the theorem,

$$\begin{aligned} 0 = \langle D_V X, V\rangle &= \langle D_V(X_3\mathcal{R} + \partial_s)\ ,\ V\rangle \\ &= \langle D_V(X_3\mathcal{R})\ ,\ V\rangle && \text{(since } \partial_s \text{ is parallel)} \\ &= X_3\ \langle D_V\mathcal{R}\ ,\ V\rangle && \text{(since } V(X_3) = 0) \\ &= X_3\ \|V\|^2 \end{aligned}$$ ∎

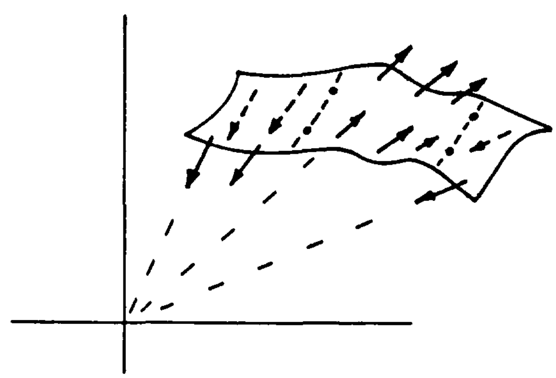


Figure 9

A second corollary of a similarly simple nature further clarifies those X which generate isometries of $\left\{\omega=0\ ,\ ds=0\right\}$.

23. COROLLARY. (uniform motion) Let X on U be admissible and let sterance g and 1-form ω both be compatible with X . Let φ_Δ be an isometry of $\left\{\omega=0\ ,\ ds=0\right\}$. If $X = hP + \partial_s$, where $D_Y P = 0$ for all Y (including $Y = \partial_s$), then h is a constant.

Proof. By the statement 4 of the theorem

$$0 = \left\langle D_{\left(\frac{a_j}{\|a_j\|}\right)}(hP + \partial_s)\ ,\ \frac{a_j}{\|a_j\|}\right\rangle$$

$$= dh\left(\frac{a_j}{\|a_j\|}\right)\ \left\langle P\ ,\ \frac{a_j}{\|a_j\|}\right\rangle\ ,\quad j = 1,2.$$

Fix j . If $dh\left(\frac{a_j}{\|a_j\|}\right) \neq 0$ at (m,t) , then there is a neighborhood in which $\left\langle P\ ,\ \frac{a_j}{\|a_j\|}\right\rangle = 0$ by smoothness of P and a_j . But then

$$0 = X\left\langle P\ ,\ \frac{a_j}{\|a_j\|}\right\rangle = 0 + \left\langle P\ ,\ D_X\left(\frac{a_j}{\|a_j\|}\right)\right\rangle = \left\langle P\ ,\ D_{\left(\frac{a_j}{\|a_j\|}\right)}X\right\rangle = dh\left(\frac{a_j}{\|a_j\|}\right)\|P\|^2\ ,$$

which is a contradiction. ∎

These two corollaries give some insight into the permissible X . They do not depend on the sterance g ; they depend only on the isometry assumption. In the next section the sterance is used to clarify a question of the uniqueness of X for a given sterance and a given 1-form.

2.5 Unique flows

In this section we make explicit use of a_g and a_ν . We recall a few facts from Table 2 (in Proposition 19). The vector fields a_g and a_ν are tangent to the integral manifolds of $\{\omega=0\ ,\ ds=0\}$; $\langle a_g\ ,\ a_\nu\rangle = 0$; and

$$dg(a_g) = 1\ , \qquad \nu(a_g) = 0\ ,$$
$$dg(a_\nu) = 0\ , \qquad \nu(a_\nu) = 1\ .$$

That is, since $dg(a_\nu) = 0$, a_ν is tangent to the level sets of g as well as tangent to the integral manifold. This holds, by construction, for any 1-form ω used to define the integral manifold. On the other hand, a_g is tangent to the integral manifold but is "adjusted" so that $\langle a_g\ ,\ a_\nu\rangle = 0$. In Figure 10 a sphere is illustrated in which the level sets of g are lines of latitude.

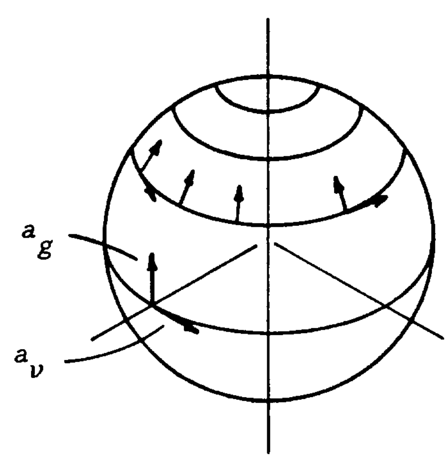


Figure 10

The next result is essentially a continuation of Theorem 21.

24. THEOREM. Let X be an admissible vector field on U, with a sterance g, $dg - \frac{\partial g}{\partial s}ds \neq 0$, and 1-form ω compatible with X. Let $\{a_i\}$ be the vector fields dual to $a^{1*}=dg$, $a^{2*}=\nu$, $a^{3*}=\omega$, $a^{4*}=ds$. With the notation $a_1=a_g$ and $a_2=a_\nu$, if $a_\nu\left(\|a_g\|^2\right) \neq 0$, then the following are equivalent.

1. The local one parameter group of X, where defined, is an isometry of $\left\{\omega=0 , ds=0\right\}$.

2. $$X\langle a_g , a_g\rangle = X\left(\|a_g\|^2\right) = 0$$
$$X\left[a_g\left(\|a_g\|^2\right)\right] = 0$$
$$X\left[\frac{a_\nu}{\|a_\nu\|}\left(\|a_g\|^2\right)\right] = 0$$

3. For any sequence of vector fields $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots, \mathcal{A}_d$, where each $\mathcal{A}_j$ is either a_g or $\frac{a_\nu}{\|a_\nu\|}$, the function $\mathcal{A}_d\left(\mathcal{A}_{d-1}\left(\ldots\mathcal{A}_1\left(\|a_g\|^2\right)\ldots\right)\right)$ is annihilated by X.

Proof. $1 \Longleftrightarrow 2$. The forward ($\Rightarrow$) implication clearly holds, for, by Theorem 21, statement 3, X commutes with both a_g and $\frac{a_\nu}{\|a_\nu\|}$, and $X\left(\|a_g\|^2\right) = 0$.

The converse requires the additional assumption in the statement of the theorem. For in general $[X, a_g] = \nu\left([X, a_g]\right)a_\nu$ and $\left[X , \frac{a_\nu}{\|a_\nu\|}\right] =$

$\nu\left(\left[X , \frac{a_\nu}{\| a_\nu \|}\right]\right) a_\nu$. But by the conditions in 2

$$\nu\left([X, a_g]\right) a_\nu\left(\| a_g \|^2\right) = [X, a_g]\left(\| a_g \|^2\right) = 0 ,$$

$$\nu\left(\left[X , \frac{a_\nu}{\| a_\nu \|}\right]\right) a_\nu\left(\| a_g \|^2\right) = \left[X , \frac{a_\nu}{\| a_\nu \|}\right]\left(\| a_g \|^2\right) = 0 .$$

If $a_\nu\left(\| a_g \|^2\right) \neq 0$, then these two equations provide the result.

Statement 3 clearly implies 2, and it follows from 1 by using the commutativity of X with both a_g and $\frac{a_\nu}{\| a_\nu \|}$, and $X\left(\| a_g \|^2\right) = 0$. ∎

The significance of the theorem is that it provides relationships between the vector field X and the 1-form ω, for a_g and a_ν are defined in terms of ω . (Recall, e_g and e_β are determined by g alone.) The theorem is an example of how we are presently seeking to understand the sense in which the pair X and ω can be determined by g under the restriction that X generate an isometry of $\left\{\omega=0 \ , \ ds=0\right\}$.

The following theorem explains the significance of the condition that $a_\nu\left(\| a_g \|^2\right) \neq 0$. In fact, if $a_\nu\left(\| a_g \|^2\right) \neq 0$, then there is a uniqueness result that uses somewhat less than the condition that X generate an isometry.

25. THEOREM. Let X be an admissible vector field on $U \subset (\mathbb{R}^3-\{0\})\times I$, and let g, $dg - \frac{\partial g}{\partial s}ds \neq 0$, be a sterance and ω a 1-form both of which are compatible with X. Let $\{a_i\}$ be the vector fields dual to $a^{1*}=dg$, $a^{2*}=\nu$, $a^{3*}=\omega$, $a^{4*}=ds$, with the notation $a_1=a_g$ and $a_2=a_\nu$. If $X\left(\|a_g\|^2\right) = 0$ (for example, if X generates an isometry of $\left\{\omega=0\ ,\ ds=0\right\}$), then X is unique whenever $a_\nu\left(\|a_g\|^2\right) \neq 0$ on U.

Proof. Since $a^{i*}(a_j) = \delta_{ij}$, the four 1-forms dg, $d\left(\|a_g\|^2\right)$, ω, ds are linearly independent if and only if $a_\nu\left(\|a_g\|^2\right) \neq 0$. Since $\{a^{i*}\}_i$ and $\{a_i\}_i$ are determined by g and ω, then the four vector fields dual to dg, $d\left(\|a_g\|^2\right)$, ω, ds are determined by g and ω. In particular, X is so determined since $dg\ X = d\left(\|a_g\|^2\right)\ X = \omega(X) = 0$ and $ds(X) = 1$. ■

26. COROLLARY. With the conditions as in the theorem, since $\omega = dF$ locally, the functions $y_1 = g$, $y_2 = \|a_g\|^2$, $y_3 = F$, $y_4 = s$ are a coordinate system for $U \subset (\mathbb{R}^3-\{0\})\times I$. For this coordinate system $X = \frac{\partial}{\partial y_4}$, the slices $\left\{y_1\ ,\ y_2\ ,\ y_3=\text{const}\ ,\ y_4=\text{const}\right\}$ are the integral manifolds of $\left\{\omega=0\ ,\ ds=0\right\}$, and $\frac{\partial}{\partial y_1}$, $\frac{\partial}{\partial y_2}$ are tangent to $\left\{\omega=0\ ,\ ds=0\right\}$.

This theorem can be compared to our result in Chapter 1, Sec. 4. There we had a uniquely determined sequence of submanifolds of $\mathbb{R}^3$, but no canonical way to define a flow that generated this sequence. This

theorem says there is at most one flow that preserves $\|a_g\|^2$, if $a_\nu\left(\|a_g\|^2\right) \neq 0$.

It is easy to see that the condition $a_\nu\left(\|a_g\|^2\right) \neq 0$ is necessary. In fact, a proof by picture is already available in Figure 10. If the level sets of g correspond to lines of latitude on the sphere, then $a_\nu\left(\|a_g\|^2\right) = 0$. Any rotation of the sphere about the north, south axis is compatible with g and with the 1-form ω that characterizes a stationary sphere.

PART 2

SOME SOLUTION TO PROBLEMS IN VISION FROM DECONVOLUTION METHODS

3 DECONVOLUTION FOR THE CASE OF MULTIPLE CHARACTERISTIC FUNCTIONS OF CUBES IN R^n

Summary

Explicit error bounds are exhibited for a case of deconvolution with elementary convolutors on $\mathbb{R}^n$. The convolutors studied are a set of $n+1$ characteristic functions of cubes (e.g., with side length $\sqrt{j}$, $j=1,2,\ldots n+1$) which operate by convolution on $L^1 \cap L^2(\mathbb{R}^n)$. For a suitable choice of approximate identity, a set of $n+1$ functions (deconvolutors) in $L^2(\mathbb{R}^n)$ are exhibited which restore $L^1 \cap L^2(\mathbb{R}^n)$, up to convolution with the approximate identity, from the $n+1$ convolutions. For the case of the convolutors operating on $L^1 \cap L^2 \cap L^p(\mathbb{R}^n)$, $1 \le p < \infty$, explicit bounds for the restoration error in the norm $L^p(E)$, E compact, are exhibited; that is, error bounds for restoration restricted to a

compact subset. The motivation for this study is the digital implementation of this deconvolution for the application to signal detectors which act by integrating over cubic regions. This motivation is discussed along with remarks on the significance of the topology for signals that is implied by the notion of restoration or deconvolution.

3.1 INTRODUCTION: DECONVOLUTION AND MACHINE VISION

Our interest in deconvolution is in part a consequence of a point of view in machine vision that we have been developing. (For contemporary developments in machine vision see, for example, Marr (1982).) In this introduction we shall indicate this point of view and we shall also indicate certain constraints to deconvolution that arise in machine vision.

The deconvolution problems that are of interest here are of the type: on $\mathbb{R}^n$, given N distributions of compact support $\mu_1, \mu_2, \ldots, \mu_N$ (called convolutors), determine the existence, support, and construction of N distributions $\nu_1, \nu_2, \ldots, \nu_N$ (called deconvolutors) such that

$$\sum_{i=1}^{N} \mu_i * \nu_i = \delta ,$$

where δ is the Dirac distribution.

For machine vision the interest is in $\mathbb{R}^2$ or $\mathbb{R}^3$. Existence of the deconvolutors depends on the μ_i: e.g., the μ_i cannot all be smooth (C^∞) functions. A condition can be placed on the μ_i, called strong coprimeness, such that the desired ν_i exist and have compact support (Kelleher and Taylor 1971). The cases for which the μ_i are characteristic functions of *a*) two intervals on $\mathbb{R}$ and *b*) two disks on $\mathbb{R}^2$ have been examined in Berenstein and Yger (1983) and in Berenstein, Taylor, and Yger (1983a, 1983b). For these cases deconvolutors with

compact support exist when, for example, two interval lengths or two disk diameters have the ratios $\sqrt{2}$ or 2, respectively. Explicit formulas for the deconvolutors in cases *a*) are reported in Berenstein, Krishnaprasad, and Taylor (1984).

Let us consider a role for deconvolution, or signal reconstruction, in machine vision. In machine vision one seeks information about objects by means of one or more images. Let us consider objects that can be modeled as a finite union $\bigcup_j M_j$ of C^1 2-manifolds M_j in $\mathbb{R}^3$. An emitted or reflected radiation can be associated with an object by defining a density F on the sphere bundle of $\mathbb{R}^3$ restricted to $\bigcup_j M_j$, $S\mathbb{R}^3|_{\bigcup_j M_j}$, where the density is with respect to a choice of volume form for $S\mathbb{R}^3|_{\bigcup_j M_j}$. To include the variable time we consider the product space $S\mathbb{R}^3|_{\bigcup_j M_j}\times\mathbb{R}$. Let $\mathcal{M}$ denote a subset of the set of such densities along with their support

$$\mathcal{M} \subset \{F : S\mathbb{R}^3|_{\bigcup_j M_j}\times\mathbb{R} \longrightarrow \mathbb{R}\}.$$

Let E_2 denote a subset of $\mathbb{R}^2$. This subset will represent what is typically referred to as the "image plane." Let $\mathcal{F}$ denote a subset of the set of time varying image densities

$$\mathcal{F} \subset \{f \ ; \ E_2\times\mathbb{R} \longrightarrow \mathbb{R}\}.$$

A basic problem in machine vision is the definition and construction of a suitable left inverse ρ of an image forming map p,

$$\mathcal{M} \underset{\rho}{\overset{p}{\rightleftarrows}} \mathcal{F} .$$

Additionally, and most importantly, appropriate topologies are sought

so that ρ is continuous. For example, if $\mathcal{M}$ is a finite set with the discrete topology, then $\mathcal{F}$ should consist of disconnected components, each containing at most one point from $p(\mathcal{M})$, and on each component ρ is constant. If $p(m)$ is in component C then the convolution $\varphi * p(m)$ of a given function φ with $p(m)$ may not be in C. In this example, the role for deconvolution is to map $\varphi * p(m)$ back to C. Since we require only that the deconvolution yield a point in a neighborhood of $p(m)$, we use the term approximate deconvolution.

We shall leave further mathematical details on this point of view to a future paper, but we will discuss the motivation. The motivation is that we wish to consider separately the questions of image quality and the questions of machine vision, and then to join these questions throught continuity of vision on an appropriate space of images. We separately consider these questions because it seems ill advised to address the issue of vision over some neighborhood of an image (which might include the image plus some additive noise, convolutions of the image, or non-linear sensor degradation of the image) when the issue of vision at the idealized, perfect image remains an open question. With this separation, we consider the idealized, perfect image (e.g., $p(m)$ for $m \in \mathcal{M}$) as a limit point in an appropriate function space $\mathcal{F}$, and we shall require that any well defined vision algorithm ρ be continuous on this space. (An example of a topology for $\mathcal{M}$ is the smallest topology such that p is continuous.)

We now turn to the specific issues in image quality and convolution-deconvolution that are the subject of this paper. For any image $f \in \mathcal{F}$ we never know f: we measure, for example, $\int_Q f$, where Q is a neighborhood of $(0,0) \in E_2\times\mathbb{R}$, instead of $f(0,0)$. To use our continuous vision algorithm, if we cannot know f then we would like to be sufficiently close to f. Let us consider an example of what we can know about f. The set Q could be $\left(-\frac{a}{2},\frac{a}{2}\right]\times\left(-\frac{a}{2},\frac{a}{2}\right]\times(-T,0)$. That is, Q models a square detector of side length $a > 0$ centered at $0 \in E_2 \subset \mathbb{R}^2$ and which integrates over the time interval $(-T,0)$, $T>0$. Let

$$A = \left[-\frac{a}{2},\frac{a}{2}\right)\times\left[-\frac{a}{2},\frac{a}{2}\right)\times\left(0,T\right] = \{x : -x \in Q\} \equiv -Q,$$

and let χ_A be the characteristic function of A. Then

$$\int_Q f = \int f\chi_Q = \left(\chi_A * f\right)(0,0).$$

Let us model a *staring array* with a simple integration time response. A set of non-overlapping subsets which covers $E_2\times\mathbb{R}$ (up to Lebesque measure zero) is

$$\left\{Q_{p,q} = \left[-\frac{a}{2}+p_1a\ ,\ \frac{a}{2}+p_1a\right) \times \left[-\frac{a}{2}+p_2a\ ,\ \frac{a}{2}+p_2a\right) \times \left[(q-1)T\ ,\ qT\right) :\ p=(p_1,p_2)\in\mathbb{Z}^2\ ,\ q\in\mathbb{Z}\right\}.$$

Let each $Q_{p,q}$ model a square detector of side length a centered at $pa = (p_1a, p_2a)\in E_2$ which integrates over the time interval $\left[(q-1)T\ ,\ qT\right)$. Let $\left(\chi_A\right)_{[(u,s)]}$ be the shift of χ_A by (u,s),

$$\left(\chi_A\right)_{[(u,s)]}(x,t) = \chi_A(x-u,t-s).$$

With this $\left(\chi_A\right)_{[(pa,qT)]} = \chi_{Q_{p,q}}$, and

$$\int \left(\chi_Q\right)_{[(pa,qT)]} f = \left(\chi_A * f\right)(pa,qT).$$

For the staring array we do not measure $f \in \mathcal{F}$ but rather

$$\left\{ \left(\chi_A * f\right)(pa, qT) \ : \ p \in \mathbb{Z}^2 \ , \ q \in \mathbb{Z} \ , \ (pa, qT) \in E \right\}$$

where E is some bounded subset of $E_2 \times \mathbb{R}$.

An answer to the question of what can be said about f based on the measured data is that for f in a suitable choice of normed function space, these meausred values can be used to approximate $\chi_E\left(\chi_A * f\right)$ by the interpolation

$$\chi_E \sum_{p,q} \left(\chi_A * f\right)(pa, qT) \ \psi_{p,q} \ ,$$

where $\psi_{p,q}$ is a choice of interpolating function (e.g., $\psi_{p,q} = \chi_Q(x - pa, t - qT)$). Moreover, for suitable normed spaces, $\chi_E\left(\chi_A * f\right)$ approximates $\chi_E f$. For a choice of A let $\mathcal{I}$ denote the set of all interpolations

$$\mathcal{I} = \left\{ \chi_E \sum_{p,q} \left(\chi_A * f\right)(pa, qT) \ \psi_{p,q} \ : \ f \in \mathcal{F} \right\}.$$

Let $\mathcal{I}^N$ denote the direct product of N such sets, in each of which a different characteristic function is used,

$$\mathcal{I}^N = \mathop{\times}_{i=1}^{N} \left\{ \chi_E \sum_{p,q} \left(\chi_{A_i} * f\right)(pa_i, qT) \ \psi_{p,q,i} \ : \ f \in \mathcal{F} \right\}.$$

Thus, what is known about f is a set of approximating interpolations of approximating convolutions.

We summarize the above by the diagram

$$\begin{array}{ccc} \mathcal{M} & \overset{p}{\underset{\mathscr{p}}{\rightleftarrows}} & \mathcal{F} \\ & & r \Big\downarrow \Big\uparrow \mathscr{r} \\ & & \mathcal{I}^N \end{array} \ .$$

The relations ρ and p are between objects and images in the sense we have modeled them above. The map r is the composition

$$\text{interpolation} \circ \text{sample} \circ \text{convolution}$$

just discussed.

The map $\mathscr{r}$ is the subject of this paper. For a given choice of norm on $\mathscr{F}$, $\mathscr{r}$ is a map from $r(f)$, $f\in\mathscr{F}$, to an approximate reconstruction of f. This may be viewed as a numerical implementation of the deconvolution from $\left\{\chi_{A_i} * f\right\}_{i=1,2,\ldots,N}$ to f, for the reconstruction is based on a finite set of values from the convolutions. The existence and continuity of the operator which deconvolves $\left\{\chi_{A_i} * f\right\}_{i=1,2,\ldots,N}$ is discussed later. Given this operator, its continuity permits us to discuss approximate deconvolution based on approximations of $\chi_{A_i} * f$ by interpolation.

We now turn to a second item, certain physical constraints on deconvolution. Let A, Q, $A_{p,q}$, E, and $f\colon E_2\times\mathbb{R} \longrightarrow \mathbb{R}$ be as above.

It has already been suggested that the set $\{A_{p,q}\}$ is to be a cover of $E_2\times\mathbb{R}$ by non-overlapping sets. Recall

$$\left(\chi_A * f\right)(pa, qT) = \int \chi_{Q_{p,q}} f .$$

The physical constraint is that $Q_{p,q} \cap Q_{p',q'} = \emptyset$ for $(p,q) \neq (p',q')$. This is because two detectors cannot occupy the same space similtaneously. This constraint can be modified (e.g., using beam splitters) such that the constraint is

$$\sum_{p,q} c(p,q)\chi_{Q_{p,q}}(x,t) = 1$$

where $c:\mathbb{Z}^2\times\mathbb{Z} \longrightarrow [0,1]$. For the staring array example, $c = 1$. (Here we do not include detector efficiency in our discussion.) This constraint will determine in part the "observation points," that is, points at which $\chi_A * f$ can be evaluated. For example, the set

$$\left\{(\chi_A * f)(p\beta, t_0) \;:\; p \in \mathbb{Z}^2\right\}$$

is not physically realizable for $\beta < a$.

In addition to constraints on the points at which $\chi_E(\chi_A * f)$ is measured, we also have bounds on the measured values. From Hölder's inequality

$$\|\chi_E(\chi_A * f)\|_\infty \le \|\chi_A\|_p \; \|f\|_{p'} \;,\; \frac{1}{p} + \frac{1}{p'} = 1 \;,\; 1 \le p \le \infty \;.$$

Let $|A|$ be the three dimensional Lebesgue measure of A, i.e., $|A| = \|\chi_A\|_1 = a^2T$. Thus

$$\|\chi_E(\chi_A * f)\|_\infty \le |A|^{1/p}\|f\|_{p'} \;,$$

and for $p < \infty$, $f \in L^{p'}(\mathbb{R}^n)$,

$$\|\chi_E(\chi_A * f)\|_\infty \xrightarrow[|A|\longrightarrow 0]{} 0.$$

(For $p = \infty$ and for $f \in L^1(\mathbb{R}^n)$, we get pointwise convergence by the Dominated Convergence Theorem, $[\chi_E(\chi_A * f)](x,t) \xrightarrow[|A|\longrightarrow 0]{} 0$.)

In the case where noise or errors for each measurement do not decrease as $|A|^{1/p}$, and for f with unit $L^{p'}$ norm, we have a lower bound α_0 on $|A|$ imposed as an addition constraint.

For A as above, the simplest pattern of observation points in $\mathbb{R}^2\times\mathbb{R}$ is the staring array with simple integrator,

$$\{(pa, qT) \;:\; p \in \mathbb{Z}^2 \;,\; q \in \mathbb{Z}\}.$$

See Figure 11. With $|A| = \alpha_0$, we may rescale a and T,

$$A(s) = \left(-s\frac{a}{2}\ ,\ s\frac{a}{2}\right) \times \left(-s\frac{a}{2}\ ,\ s\frac{a}{2}\right) \times \left[0\ ,\ {}^{T}/_{s^2}\right)\ ,\ s>0\ ,$$

so that $|A(s)| = |A| = \alpha_0$. See Figure 12*a* and 12*b*. In other words, the detector size can be reduced if the integration time is appropriately increased, and visa versa, without altering the upper bound due to Hölder.

A second simple observation scheme is a *continuous scanning* pattern. Let v be a unit vector in $\mathbb{R}^2$ and define

$$B = \left\{(x,t) \in \mathbb{R}^2\times[0,T]\ :\ x-vt \in \left(-\frac{a}{2},\frac{a}{2}\right)\times\left(-\frac{a}{2},\frac{a}{2}\right) \subset \mathbb{R}^2\right\}.$$

See Figure 13. Note that $|B| = |A|$.

A third scheme is an alternative to the continuous scan, the *shift scanning* pattern of Figure 14.

In all of these cases, the number of observation points in a fixed set $E \subset E_2\times\mathbb{R}$ is approximately $|E|/\alpha_0$. We have here the "mesh size" or sampling interval bounded below due to α_0.

Let us examine the consequences of f being independent of time. Let $f(x,t) = g(x)$, and let $P(1/n) = \left(-\frac{a}{2n},\frac{a}{2n}\right) \times \left(-\frac{a}{2n},\frac{a}{2n}\right)$, $P = P(1)$. For the rescaling case of the staring array

$$\begin{aligned}\left(\chi_{A(1/n)}*f\right)(x,t) &= \left[\left(\chi_{P(1/n)}\ \chi_{[(0,n^2T)]}\right)*f\right](x,t)\\ &= \int_{[t-n^2T,t]} \int\chi_{P(1/n)}(x-y)\ g(y)\ dyds\\ &= n^2T\ \left(\chi_{P(1/n)}*g\right)(x)\\ &= a^2T\ \left(\frac{\chi_{P(1/n)}}{\|\chi_{P(1/n)}\|_1} * g\right)(x).\end{aligned}$$

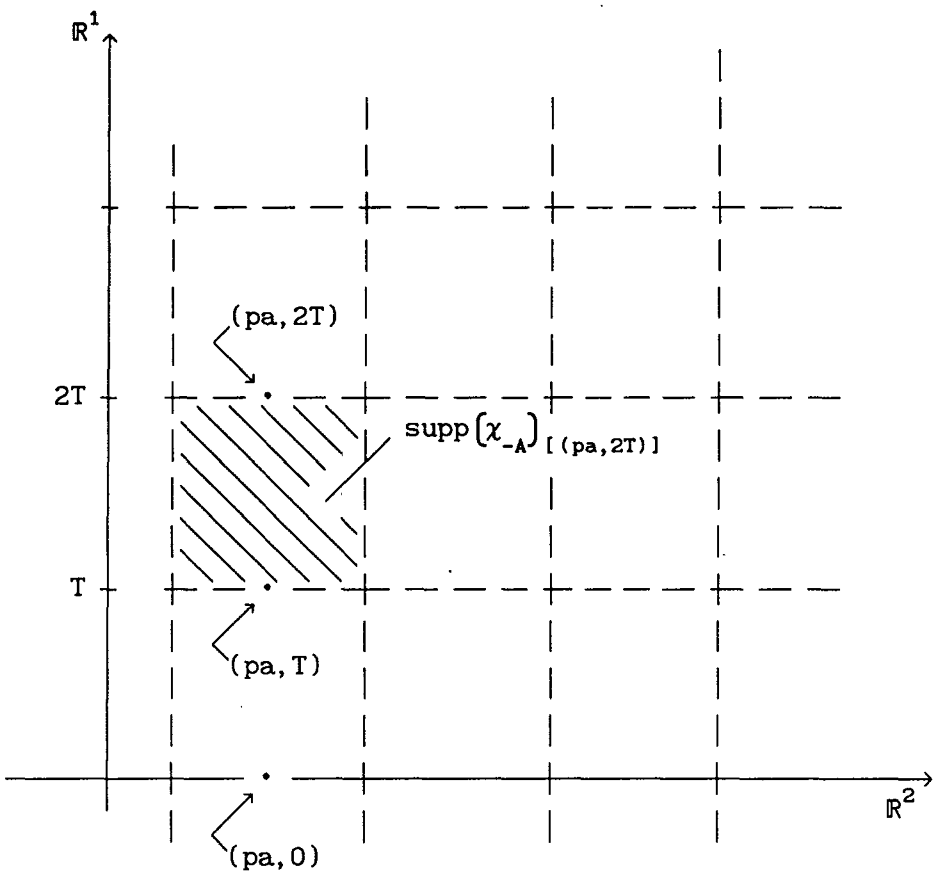


Fig. 11. Representation of a staring array with a simple integration time-response.

$$-A = \left[-\frac{a}{2},\frac{a}{2}\right)\times\left[-\frac{a}{2},\frac{a}{2}\right)\times(-T,0) \subset \mathbb{R}^3 \quad , \; p = (p_1,p_2) \in \mathbb{Z}^2$$

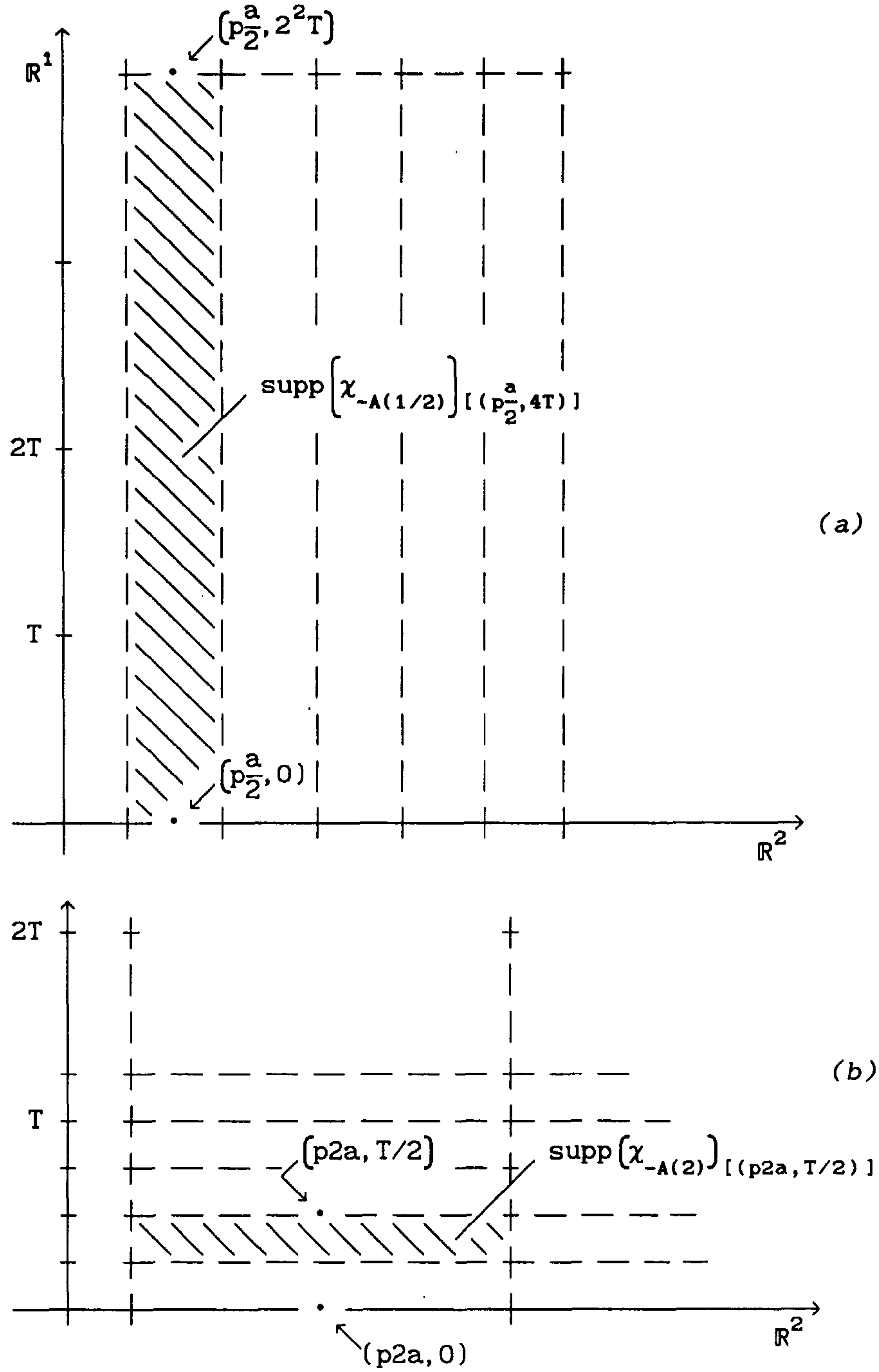


Fig. 12. Two rescalings of a and T

$|A(1/2)| = |A(2)|$

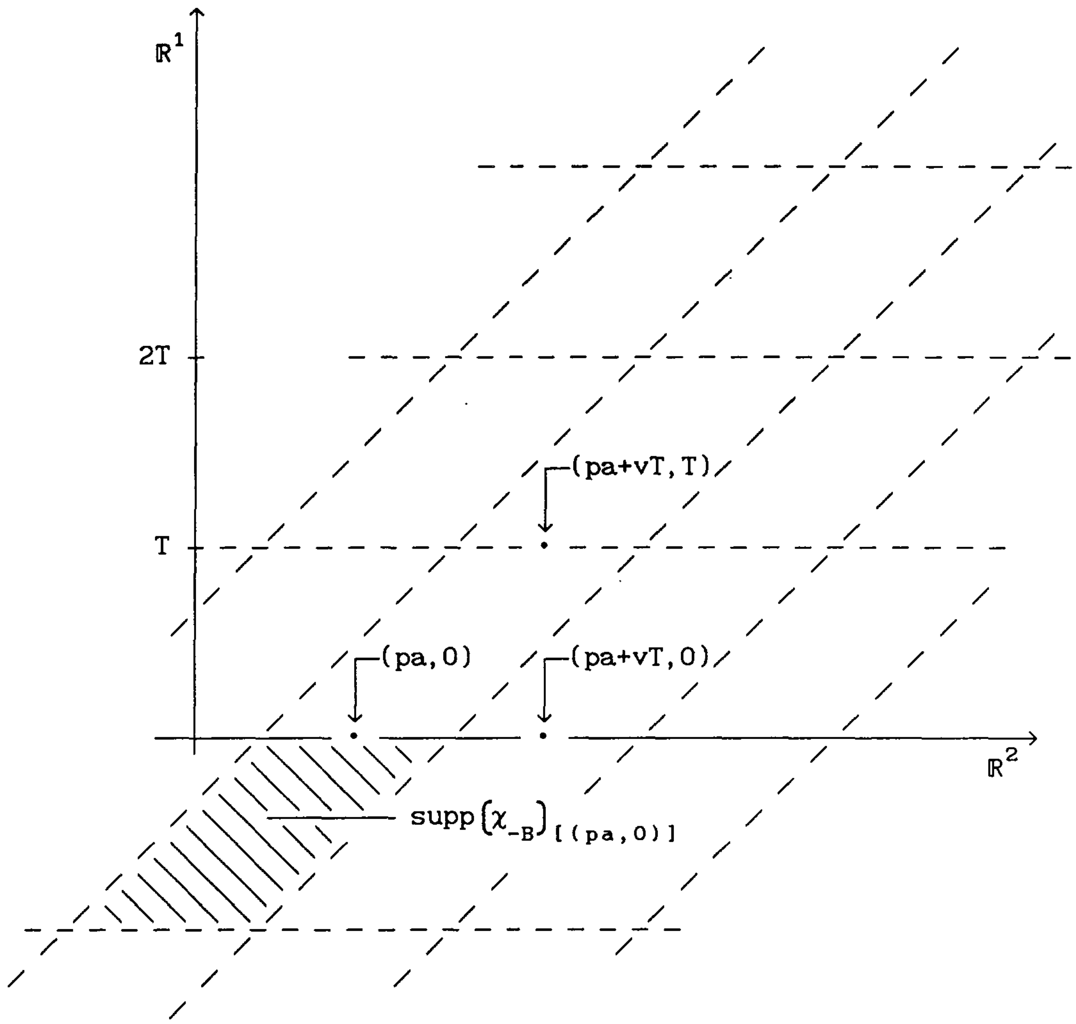


Fig. 13. Representation of continuous scanning with scan velocity $v \in \mathbb{R}^2$. In this representation

$$\left\{pa+vT \ : \ p \in \mathbb{Z}^2\right\} = \left\{pa \ : \ p \in \mathbb{Z}^2\right\} .$$

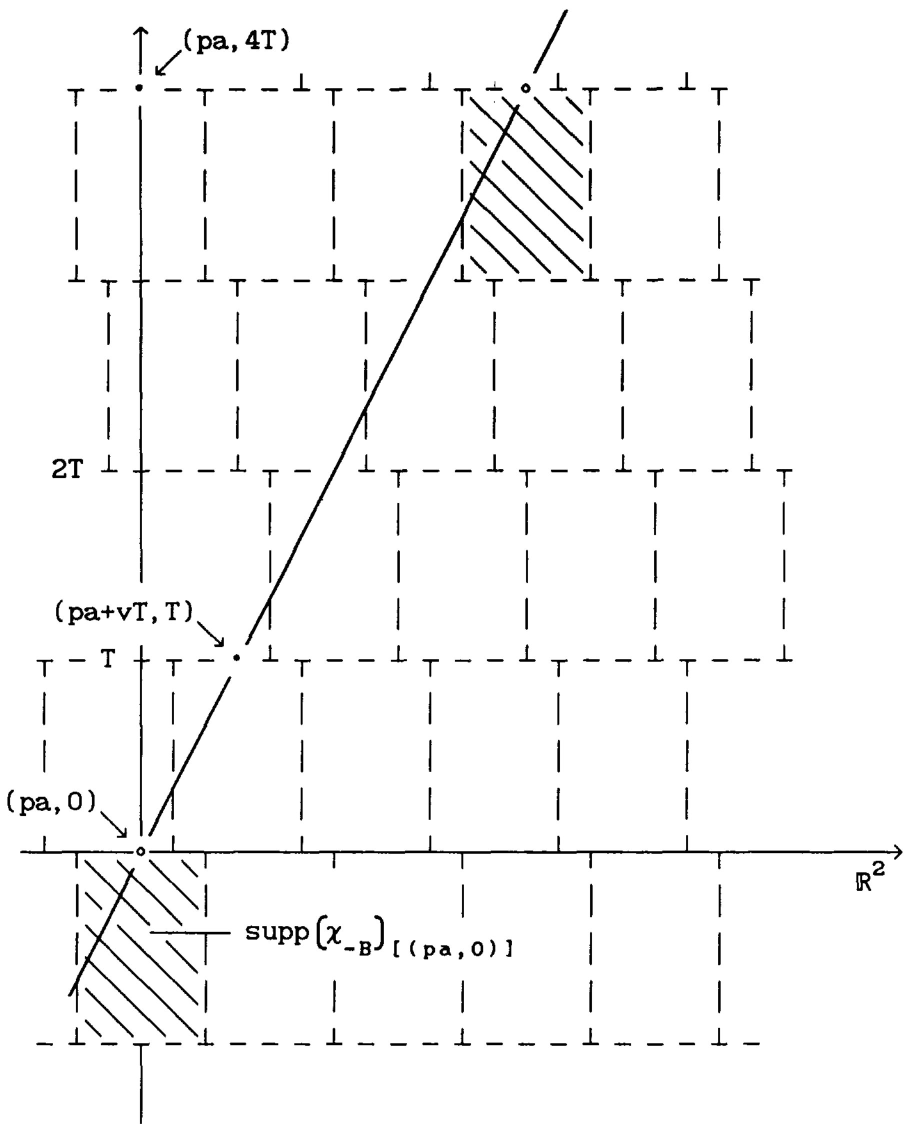


Fig. 14. Representation of shift scanning. After each interval T the set B is shifted by vT, $v \in \mathbb{R}^2$. Here, for each $n \in \mathbb{Z}$, $\{pa+nvT : p \in \mathbb{Z}^2\} = \{pa+(n+4)vT : p \in \mathbb{Z}^2\}$.

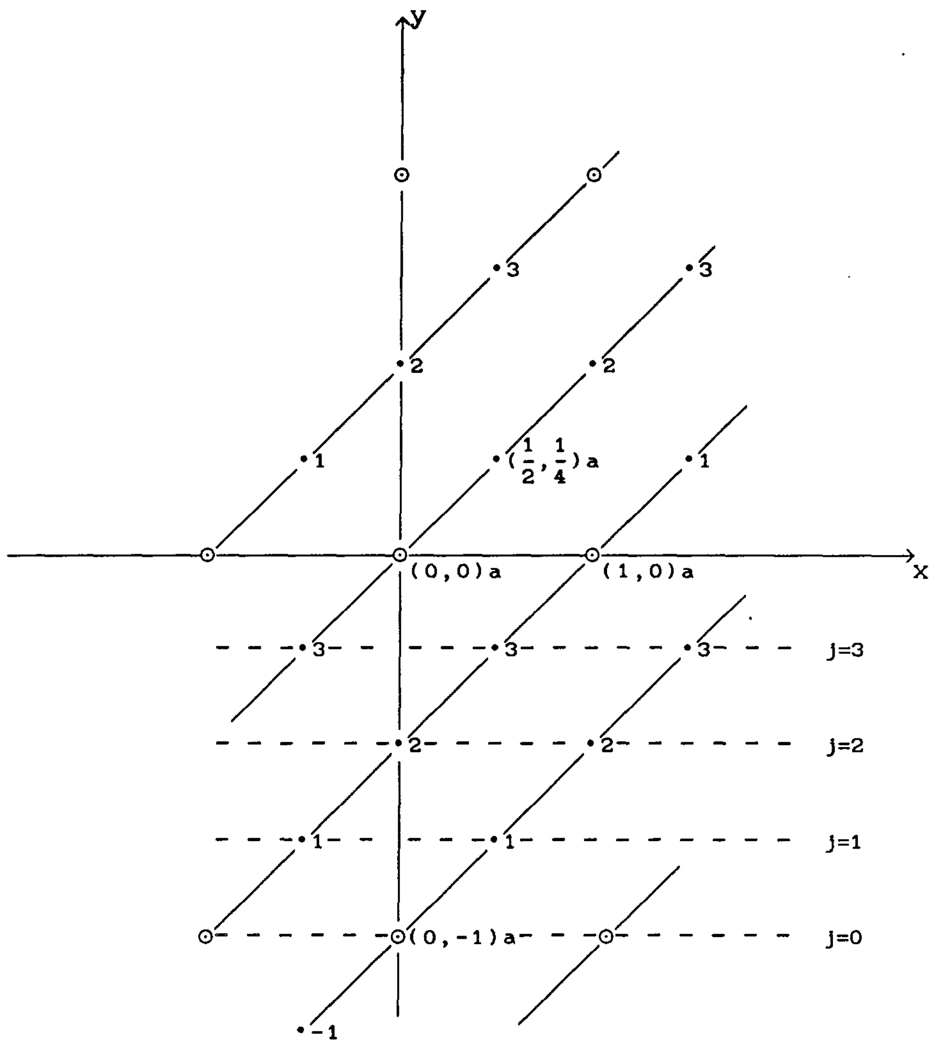


Fig. 15. Scan in $\mathbb{R}^2$ for $(p_1, p_2)+j\left(\frac{1}{2}, \frac{1}{2^2}\right)$, $(p_1, p_2) \in \mathbb{Z}^2$, $j \in \mathbb{Z}$.

Numbers at points refer to value of j for which sampled.

The observation points are $(p\frac{a}{n}, qn^2T)$, $p\in\mathbb{Z}^2$, $q\in\mathbb{Z}$, and

$$\left(\chi_{A(1/n)}*f\right)\left(p\frac{a}{n}, qn^2T\right) = a^2T\left(\frac{\chi_{P(1/n)}}{\|\chi_{P(1/n)}\|_1} * g\right)\left(p\frac{a}{n}\right).$$

We conclude

Remark 1 For rescaling of A to A(1/n), and for $g \in L^p(\mathbb{R}^n)$, $1\le p<\infty$,

a) The underlying convolution which is sampled converges in L^p to g as n increases:

$$\alpha_0\left[\frac{\chi_{P(1/n)}}{\|\chi_{P(1/n)}\|_1} * g\right] \xrightarrow[n\to\infty]{} \alpha_0 g \quad \text{in } L^p(\mathbb{R}^n) \ ;$$

b) the number of observation points on any set E_2 increases as n^2;

c) the time interval associated with each observation point increases in length as n^2.

For shifted scanning, a choice for an observation point set is

$$\left\{\left((p_1,p_2)a + j(\tfrac{1}{n}, (\tfrac{1}{n})^2)a \ , \ jT\right) \ : \ (p_1,p_2)\in\mathbb{Z}^2 \ , \ j\in\mathbb{Z}\right\},$$

whereas the set A remains the same for any n. See Figure 15. Moreover,

$$\begin{aligned}&\left\{\left(\chi_A*f\right)\left(pa + ja(\tfrac{1}{n}, (\tfrac{1}{n})^2) \ , \ jT\right) \ : \ p\in\mathbb{Z}^2 \ , \ j\in\mathbb{Z}\right\}\\ &= \left\{T\left(\chi_P*g\right)\left(pa + ja(\tfrac{1}{n}, (\tfrac{1}{n})^2)\right) \ : \ p\in\mathbb{Z}^2 \ , \ j\in\mathbb{Z}\right\}\\ &= \left\{T\left(\chi_P*g\right)\left(k(0,1)\tfrac{a}{n} + j(1,\tfrac{1}{n})\tfrac{a}{n}\right) \ : \ (j,k) \in \mathbb{Z}^2\right\} .\end{aligned}$$

We conclude

Remark 2 For shifted scanning according to

$$\left\{\left(pa + ja(\tfrac{1}{n}, (\tfrac{1}{n})^2) \ , \ jT\right) \ : \ p\in\mathbb{Z}^2 \ , \ j\in\mathbb{Z}\right\}$$

a) the underlying convolution remains

$$\alpha_0 \frac{\chi_p * g}{\|\chi_p\|_1};$$

b) the number of observation points on any set increases as n^2;

c) the time required to acquire a full observation set increases as n^2.

Let us compare rescaling (Remark 1) and scanning (Remark 2). For α_0 fixed and for f independent of time, to reduce the mesh size of the observation points projected onto E_2, we can use smaller detectors and observe over a longer time interval, or we can use (shifted) scanning and a sequence of time intervals. In both cases $|A| = \alpha_0$ and the total observation time to get all observation points is the same. Therefore,

1. Rescaling and scanning are equivalent in terms of observation time required. However,

2. Rescaling and scanning differ in that

   a) rescaling uses decreasing detector size to approach the desired function g,

   b) scanning uses a fixed detector size to approach $\chi_p * g$.

Our interest is in the scanning case. In particular, we examine the case of

$$\bigtimes_{i=1}^{3} \left\{ \left(\chi_{p_i} * g\right)(x_j) \;:\; j \in \mathbb{Z} \right\},$$

that is, more than one detector of fixed but appropriate sizes and a sequence of observation points whose mesh size can be as fine as required. For such a case, the desired approximate reconstruction of g can be given.

3.2 Construction of an approximate deconvolution on $\mathbb{R}^n$

While we shall address in detail the case in which the convolutors are the characteristic functions of cubes in $\mathbb{R}^n$ (e.g., the cubes P_1, P_2, and P_3 in $\mathbb{R}^2$ mentioned just above), we may begin somewhat more generally. Specifically, we shall assume we are given N convolutors μ_i, $i=1,2,\dots N$, and each μ_i is in $L^\infty(\mathbb{R}^n)$ and has compact support. Let f be in $L^1 \cap L^\infty(\mathbb{R}^n)$. We wish to approximately reconstruct f from $r(f) \in \mathcal{I}^N$, where r and $\mathcal{I}^N$ are as in the Introduction of this chapter. Approximate will mean any of the L^p norms, $1 \le p < \infty$ (and $p=\infty$ with some additional qualifications).

For approximate reconstruction of f it suffices that for sufficiently large τ, $\tau>0$, the reconstruction approximate $\varphi_\tau * f$, where $\varphi \in L^1(\mathbb{R}^n)$ and $\varphi_\tau(x) = \tau^n \varphi(\tau x)$ for $x \in \mathbb{R}^n$, since

$$\varphi_\tau * f \xrightarrow[\tau \to \infty]{} f \quad \text{in } L^p(\mathbb{R}^n),\ 1 \le p < \infty .$$

In this case we seek N deconvolutors $\nu_1, \nu_2, \dots, \nu_N$ such that

$$\sum_{i=1}^{N} \mu_i * \big(\nu_i * \varphi_\tau\big) * f = \varphi_\tau * f.$$

The ingredients for a solution $\{\nu_i * \varphi_\tau\}_{i=1,2,\dots,N}$ were noted by Berenstein, Krishnaprasad, and Taylor (1984). Let $^\wedge$ denote the Fourier transform. The Fourier transform of distributions in the equation $\sum_{i=1}^{N} \mu_i * \nu_i = \delta$ results in the Bezout equation $\sum_{i=1}^{N} \hat{\mu}_i\ \hat{\nu}_i = 1$. A necessary condition on $\{\mu_i\}_{i=1,\dots,N}$ for the existence of a solution is thus

$\sum_{i=1}^{N} |\hat{\mu}_i|^2(\omega) > 0$ for all $\omega \in \mathbb{R}^n$. For such μ_i a solution of $\sum_{i=1}^{N} \hat{\mu}_i \; D_i = 1$ is

$$D_i = \frac{\overline{\hat{\mu}_i}}{\sum_{i=1}^{N} |\hat{\mu}_i|^2},$$

where $^-$ denotes complex conjugation. However, the D_i are not the solutons $\hat{\nu}_i$ if each ν_i is to be a distribution with compact support, for the D_i are not analytic. On the other hand we have the following. For $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ let

$$\|\omega\|_\infty = \max_{j=1,2,\ldots,n} \left\{ |\omega_j| \right\}.$$

The growth of D_i as $\|\omega\|_\infty$ gets large is known once a lower bound is established for $\sum_{i=1}^{N} |\hat{\mu}_i|^2(\omega)$. For the μ_i of interest we shall exhibit such a bound as well as a choice of φ_τ such that $D_i \hat{\varphi}_\tau \in L^2(\mathbb{R}^n)$. In this case there exists $h_i \in L^2(\mathbb{R}^n)$ such that $\hat{h}_i = D_i \hat{\varphi}_\tau$ and

$$\sum_{i=1}^{N} h_i * \mu_i * f = \varphi_\tau * f,$$

this last equation easily seen by taking Fourier transforms. (We have assumed $f \in L^1(\mathbb{R}^n)$ so that $\mu_i * f \in L^1(\mathbb{R}^n)$ and the left hand side is in $L^2(\mathbb{R}^n)$.)

The $\{h_i\}_{i=1,\ldots,N}$ are the desired approximate deconvolutors. However, they do not have compact support. On the other hand, they can be explicitly exhibited using only the knowledge of the Fourier transforms of the convolutors μ_i. Because of this simplicity and potential utility, in the following we conduct an error analysis for a digital implementation of this approximate deconvolution for the

special case of convolutors which are characteristic functions of cubes in $\mathbb{R}^n$. In addition the cases described in the introduction provide two further restrictions on the problem and these we adopt.

First, it will suffice to approximately reconstruct f on some compact set E. For example, it suffices to choose τ such that

$$\varepsilon_1 = \|\chi_E(f - \varphi_\tau * f)\|_p$$

is sufficiently small.

Second, the measurements consist of a discrete set in $\mathbb{R}^n$ on which a set of convolutions is evaluated. Let $\{x_q\}_{q\in Q}$ denote the discrete set of points, with $x_q \in \mathbb{R}^n$ and with Q a finite index set. The convolution values are

$$\left\{(\mu_i * f)(x_q) \ : \ q\in Q,\ \ i=1,2,\ldots,N\right\}.$$

We seek to use these values to approximate f on E by constructing an interpolation. In particular we seek functions $\psi_j : \mathbb{R}^n \longrightarrow \mathbb{R}$, $j \in J \subset Q$, and we seek a map $\tilde{G}: \{x_j\}_{j\in J} \longrightarrow \mathbb{R}$ such that $\chi_E = \chi_E \sum_{j\in J} \psi_j$ and $\chi_E f$ is approximated in L^p by $\chi_E \sum_{j\in J} \tilde{G}(x_j)\psi_j$. That is, we seek to make the error ε ,

$$\varepsilon = \|\chi_E(f - \sum_{j\in J} \tilde{G}(x_j)\psi_j)\|_p ,$$

sufficiently small.

For brevity let F denote $\varphi_\tau * f$. From above we have

$$F \equiv \varphi_\tau * f = \sum_{i=1}^{N} h_i * \mu_i * f \ .$$

Let $^\vee$ denote inverse Fourier transform, let χ_λ be the characteristic function of the set $\{\omega\in\mathbb{R}^n : \|\omega\|_\infty \le \lambda\}$, and let $\mathcal{B}$ be a compact set in $\mathbb{R}^n$

with characteristic function $\chi_{\mathcal{B}}$. Then define

$$G = \sum_{i=1}^{N} \Big[\big(\chi_\lambda \hat{h}_i\big)^{\vee} \chi_{\mathcal{B}} \Big] * \big(\mu_i * f\big).$$

We shall seek to choose λ and $\mathcal{B}$ such that

$$\varepsilon_3 = \|\chi_E \sum_{j\in J} \big(F(x_j) - G(x_j)\big)\psi_j\|_p$$

is sufficiently small.

The triangle inequality now indicates the additional two terms needed to have a bound for ε. One term is

$$\varepsilon_2 = \|\chi_E \ \big(F - \sum_{j\in J} F(x_j)\psi_j\big)\|_p ,$$

and the second term is

$$\varepsilon_4 = \|\chi_E \sum_{j\in J} \big(G(x_j) - \tilde{G}(x_j)\big)\psi_j\|_p .$$

The defining expression for G above suggests the consideration of $\tilde{G}(x_j)$ of the form

$$\tilde{G}(x_j) = \sum_{i=1}^{N} \sum_{q\in Q} \tilde{H}_i(x_j - x_q)\big(\mu_i * f\big)(x_q) ,$$

where $\tilde{H}_i : \{x_q\}_{q\in Q} \longrightarrow \mathbb{R}$. The $\tilde{H}_i$ then are the deconvolutors which we shall implement. An L^p error bound for this approximation is thus

$$\varepsilon = \|\chi_E \ \big(f - \sum_{j\in J} \tilde{G}(x_j)\psi_j\big)\|_p \le \varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4 .$$

In the following we develop bounds for each of the four error terms.

3.3 The lower bound $C(\omega)$

For a choice of $n+1$ positive numbers $a_0, a_1, \ldots, a_n$ let A_i be the cube of side length a_i in $\mathbb{R}^n$. Let χ_{A_i} be the characteristic function of A_i and let $\mu_i = \dfrac{\chi_{A_i}}{(a_i)^n}$. For this specific case of convolutor we shall prove the following.

1. THEOREM. (Berenstein) Let $\alpha_0, \alpha_1, \ldots, \alpha_n$ be a choice of $n+1$ positive integers such that the collection is pairwise relatively prime and none is a perfect square. Let $a_i = \sqrt{\alpha_i}$, and let $M = \max\limits_{i=0,1,\ldots n} \{a_i\}$. Let μ_i be the characteristic function of the cube in $\mathbb{R}^n$ with side length a_i normalized to unit L^1 norm as above. Then

$$\sum_{i=0}^{n} |\hat{\mu}_i(\omega)|^2 \geq \left(\frac{8}{5M^4} \right)^{2n} \prod_{j=1}^{n} \frac{1}{\max\left\{\left(\frac{\pi}{M}\right)^4, |\omega_j|^4\right\}}, \quad \omega \in \mathbb{R}^n.$$

The proof will follow from several lemmas. We begin with

2. LEMMA. Let p and q both be postive integers. If p and q are relatively prime, then $\sqrt{p/q}$ is rational if and only if both p and q are perfect squares.

Proof. ($\Leftarrow$) Clear. ($\Rightarrow$) It follows from $qr^2 = ps^2$ for some relatively prime integers r and s, from a prime factor expansion of q, $q = q_1^{m_1}\cdots q_k^{m_k}$, and from the fact that p and q are relatively prime, that each q_i divides s. Hence,

$$s^2 = q_1^{2n_1}\cdots q_k^{2n_k}(s')^2 \ ,\ 2n_i \geq m_i \text{ for } i=1,2,\ldots,k.$$

If q is not a perfect square then, reordering factors, m_1 is odd, $2n_1-m_1 \geq 1$, hence, from $qr^2 = ps^2$, q_1 divides r^2 as well as s^2. This contradicts r and s being relatively prime. This along with a similar argument for p show that both p and q are perfect squares. ∎

To proceed it will be necessary to define some maps. Let $\mathcal{A} = \{a_0, a_1, \ldots, a_n\}$, $m = \min\limits_{a_i\in\mathcal{A}}\{a_i\}$. Let $x\in\mathbb{R}$. Define the maps

$$d:\ \mathbb{R}\times\mathcal{A} \longrightarrow \mathbb{Z}, \qquad\qquad r:\mathbb{R}\times\mathcal{A} \longrightarrow [-\frac{\pi}{2m}\ ,\ \frac{\pi}{2m}]$$

by

$$x = d(x,a_i)\ \frac{\pi}{a_i} + r(x,a_i), \qquad\qquad \operatorname{sgn}(x)\ r(x,a_i) \in [-\frac{\pi}{2a_i}, \frac{\pi}{2a_i}).$$

For fixed x we have the maps defined by restriction

$$d_x:\mathcal{A} \longrightarrow \mathbb{Z}, \qquad\qquad r_x:\mathcal{A} \longrightarrow [-\frac{\pi}{2m}\ ,\ \frac{\pi}{2m}]$$

$$d_x(a_i) = d(x,a_i) \qquad\qquad r_x(a_i) = r(x,a_i).$$

For each fixed $x\in\mathbb{R}$ we also define the subset $\Gamma_x \subset \mathcal{A}$ by

$$\Gamma_x = \left\{a \in \mathcal{A}\ :\ |r_x(a)| = \min_{a_i\in\mathcal{A}}\{|r_x(a_i)|\}\right\}.$$

A choice from Γ_x will be denoted γ_x. (The set Γ_x may consist of more than one element. Any element γ_x may be interpreted as an element of $\mathcal{A}$ for which some integer multiple of π/γ_x is as near x as any element

from the set $\{z\frac{\pi}{a_i} : z\in\mathbb{Z} , a_i\in\mathcal{A}\}$.)

The following simple observation will be used. Its proof follows directly from the definition of γ_x .

3. LEMMA. For every $Q \ge 0$ either

$$|r_x(\gamma_x)| \le Q$$

or

$$\text{for every } a_i \in \mathcal{A},\ |r_x(a_i)| > Q .$$

4. LEMMA. For every a_i, $a_j \in \mathcal{A}$ such that $a_i \ne a_j$, and for every $\delta \in \mathbb{Z} - \{0\}$,

$$|\sin\left(\frac{a_i}{a_j}\delta\pi\right)| \ge \frac{4}{\left(4\frac{a_i}{a_j}|\delta| + 1\right)a_j^{\,2}} .$$

Proof. There exists a nonnegative integer n and $\varepsilon \in [-1/2 , 1/2]$ such that

$$\frac{a_i}{a_j}|\delta|\pi = n\pi + \varepsilon\pi . \qquad (*)$$

From this and the properties of the sine function

$$|\sin\left(\frac{a_i}{a_j}\delta\pi\right)| = |\sin\left(\frac{a_i}{a_j}|\delta|\pi\right)| = |\sin(\varepsilon\pi)| \ge \frac{2}{\pi}|\epsilon\pi| . \qquad (**)$$

From $(*)$

$$|\varepsilon\pi| = \left|\frac{a_i}{a_j}|\delta| - n\right|\pi = \frac{\left|\left(\frac{a_i}{a_j}\right)^2\delta^2 - n^2\right|\pi^2}{\left|\frac{a_i}{a_j}|\delta| + n\right|\pi} . \qquad (***)$$

This cannot vanish, for

$$\left(\frac{a_i}{a_j}\right)^2\delta^2 - n^2 = 0 \iff \frac{a_i}{a_j} = \frac{n}{|\delta|},$$

hence $\frac{a_i}{a_j}$ is rational, which by Lemma 2 contradicts the intial assumptions for $\mathcal{A}$. This nonvanishing along with the fact that a_i^2 and a_j^2 are both integers implies $|a_i^2\delta^2 - n^2a_j^2| \geq 1$. This with (***) implies

$$|\varepsilon\pi| \geq \frac{\pi}{a_j^2\left|\frac{a_i}{a_j}|\delta| + n\right|}.$$

However, from (*) $n \leq \frac{a_i}{a_j}|\delta| + \frac{1}{2}$, hence

$$\left|\frac{a_i}{a_j}|\delta| + n\right| \leq \frac{a_i}{a_j}|\delta| + \frac{a_i}{a_j}|\delta| + \frac{1}{2},$$

consequently

$$|\varepsilon\pi| \geq \frac{\pi}{a_j^2\left|2\frac{a_i}{a_j}|\delta| + \frac{1}{2}\right|},$$

which, when substituted in (**), yields the desired inequality. ■

Recall $M = \max\limits_{a\in\mathcal{A}}\{a\}$.

5. LEMMA. Let $x \in \mathbb{R}$ be fixed. For every $\gamma_x \in \Gamma_x$, if

$$|d_x(\gamma_x)| \geq 1$$

and

$$|r_x(\gamma_x)| \leq \frac{2}{M}\,\frac{1}{\left(4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right)\gamma_x^2},$$

then for every $a_i \neq \gamma_x$

$$|\sin(a_i x)| \geq \frac{2}{\left[4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2} \geq \frac{2}{\left[4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2}.$$

Proof. Since $x = d_x(\gamma_x)\frac{\pi}{\gamma_x} + r_x(\gamma_x)$, by the Mean Value Theorem there exists $\xi \in \mathbb{R}$ such that

$$\sin(a_i x) = \sin\left[\frac{a_i}{\gamma_x}d_x(\gamma_x)\pi + a_i r_x(\gamma_x)\right] = \sin\left[\frac{a_i}{\gamma_x}d_x(\gamma_x)\pi\right] + a_i r_x(\gamma_x)\cos\xi.$$

Thus

$$|\sin(a_i x)| \geq \left|\sin\left[\frac{a_i}{\gamma_x}d_x(\gamma_x)\pi\right]\right| - |a_i r_x(\gamma_x)|$$

(apply Lemma 4)

$$\geq \frac{4}{\left[4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2} - a_i|r_x(\gamma_x)|$$

(apply the hypothesis)

$$\geq \frac{4}{\left[4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2} - \frac{2a_i}{M}\frac{2}{\left[4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2}$$

$$\geq \frac{2}{\left[4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2} \geq \frac{2}{\left[4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right]\gamma_x^2}.$$

∎

The next two lemmas will address the case $|x| \geq \pi/2M$. For this case the condition $|d_x(\gamma_x)| \geq 1$ is not vacuous.

6. LEMMA. If $|x| \geq \pi/2M$, then there exists $\gamma_x \in \Gamma_x$ such that
$$|d_x(\gamma_x)| \geq 1.$$

Proof. a) Case $x = \pi/2M$.

Since $x = \pi/M - \pi/2M$, then $d_x(M) = 1$ and $r_x(M) = -\pi/2M$. Moreover, for all $a_i \neq M$, $x < \pi/2a_i$, hence, $d_x(a_i) = 0$ and $r_x(a_i) = \pi/2M$, therefore $|r_x(M)| \leq |r_x(a_i)|$, or $M \in \Gamma_x$.

b) Case $x = -\pi/2M$.

Use an analogous argument with $d_x(M) = -1$, $r_x(M) = \pi/2M$, and for $a_i \neq M$ $r_x(a_i) = -\pi/2M$.

c) Case $|x| > \pi/2M$.

Consider any $\gamma_x \in \Gamma_x$. Since $x = d_x(\gamma_x)\frac{\pi}{\gamma_x} + r_x(\gamma_x)$ whereas $|r_x(\gamma_x)| \leq |r_x(M)| \leq \pi/2M < |x|$, then $|d_x(\gamma_x)| \geq 1$. ∎

7. LEMMA. If $|x| \geq \pi/2M$ then there exists $\gamma_x \in \mathcal{A}$ such that for every $a_i \in \mathcal{A} - \{\gamma_x\}$
$$\frac{|\sin(a_i x)|}{|a_i x|} \geq \frac{2}{5|x|^2 M^4} .$$

Proof. Fix $x \in \mathbb{R}$, $|x| \geq \pi/2M$. Lemma 6 provides a $\gamma_x \in \Gamma_x$ such that
$$x = d_x(\gamma_x)\frac{\pi}{\gamma_x} + r_x(\gamma_x) , \qquad |d_x(\gamma_x)| \geq 1 .$$
Since $|r_x(\gamma_x)| \leq \pi/2\gamma_x$, then $|x| \geq \left(|d_x(\gamma_x)| - \frac{1}{2}\right)\pi/\gamma_x$. Therefore,

$$\frac{1}{\left(4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right)} \geq \frac{1}{|x|}\,\frac{\left(|d_x(\gamma_x)| - \frac{1}{2}\right)\pi/\gamma_x}{\left(4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right)} .$$

The right hand expression is increasing in $|d_x(\gamma_x)|$ for $|d_x(\gamma_x)| \geq 1$, so we may use $|d_x(\gamma_x)| = 1$ to get a lower bound:

$$\frac{1}{\left(4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right)} \geq \frac{1}{|x|}\,\frac{\pi/2\gamma_x}{4a_i/\gamma_x + 1} \geq \frac{1}{|x|}\,\frac{\pi/2}{4M + \gamma_x} \geq \frac{1}{|x|}\,\frac{\pi/2}{5M} .$$

By Lemma 3 it suffices to consider two cases.

Case 1: $|r_x(\gamma_x)| \leq \dfrac{2}{M}\,\dfrac{1}{\left(4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right)\gamma_x^{\,2}}$.

In this case, if $a_i \neq \gamma_x$, then by Lemma 5

$$|\sin(a_i x)| \geq \frac{2}{\left(4\frac{a_i}{\gamma_x}|d_x(\gamma_x)| + 1\right)\gamma_x^{\,2}} ,$$

hence,

$$|\sin(a_i x)| \geq \frac{2}{\gamma_x^{\,2}}\,\frac{1}{|x|}\,\frac{\pi/2}{5M} \geq \frac{\pi}{|x|5M^3} \geq \frac{2}{|x|5M^3} ,$$

and

$$\frac{|\sin(a_i x)|}{|a_i x|} \geq \frac{2}{5|x|^2 a_i M^3} \geq \frac{2}{5|x|^2 M^4} .$$

Case 2: For every $a_i \in \mathcal{A}$, $|r_x(a_i)| > \dfrac{2}{M}\,\dfrac{1}{\left(4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right)\gamma_x^{\,2}}$.

Here, $x = d_x(a_i)\pi/a_i + r_x(a_i)$, $|r_x(a_i)| \leq \pi/2a_i$, hence

$$|\sin(a_i x)| = |\sin(a_i r_x(a_i))| \geq \frac{2}{\pi} |r_x(a_i)| a_i$$

$$\geq \frac{2a_i}{\pi} \frac{2}{M} \frac{1}{\left(4\frac{M}{\gamma_x}|d_x(\gamma_x)| + 1\right)\gamma_x^{\,2}} \geq \frac{2a_i}{\pi} \frac{2}{M} \frac{1}{|x|} \frac{\pi/2}{5M} \frac{1}{\gamma_x},$$

which implies the desired result (recall $\gamma_x \leq M$). ■

Remark: In Case 2 $a_i x$ is never an interger multiple of π. In Case 1 $a_i \neq \gamma_x$ is used.

The case $|x| \leq \pi/2M$ is addressed next. As usual $\sin(x)/x$ is defined to be 1 for $x = 0$.

8. LEMMA. If $|x| \leq \pi/2M$, then $\dfrac{|\sin(a_i x)|}{|a_i x|} \geq 2/\pi$ for every $a_i \in \mathcal{A}$.

Proof. $|x| \leq \pi/2M \Rightarrow$ for every $a_i \in \mathcal{A}$, $|x| \leq \pi/2a_i \Rightarrow$

$$\frac{|\sin(a_i x)|}{|a_i x|} \geq 2/\pi. \qquad ■$$

These last two results can be combined.

9. LEMMA. For every $x \in \mathbb{R}$, there exists $\gamma_x \in \mathcal{A}$ such that for every $a_i \in \mathcal{A} - \{\gamma_x\}$

$$\frac{|\sin(a_i x)|}{|a_i x|} \geq \frac{2}{5M^4} \frac{1}{\left[\max\left\{\frac{\pi}{2M}, |x|\right\}\right]^2}.$$

Proof. Since $M > 1$, $\frac{2}{\pi} \ge \frac{2}{\pi}\,\frac{4}{5\pi M^2} = \frac{2}{5M^4}\,\frac{1}{(\pi/2M)^2}$. With this we consider separately $|x| \le \pi/2M$ and $|x| > \pi/2M$ and apply the preceding lemmas. ∎

We can now conclude the theorem.

Proof of the theorem. Let $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \mathbb{R}^n$. It is readily checked that

$$\hat{\mu}_i(\omega) = \prod_{j=1}^{n} \frac{\sin(a_i\omega_j/2)}{a_i\omega_j/2}.$$

Let $\omega \in \mathbb{R}^n$ be fixed. For each coordinate ω_j of ω let γ_{ω_j} be an element of $\mathcal{A}$ provided by Lemma 9. Since there are at most n distinct γ_{ω_j} and since there are $n+1$ distinct elements in $\mathcal{A}$, there exists an element $a(\omega) \in \mathcal{A}$ such that $a(\omega) \ne \gamma_{\omega_j}$, $j=1,2,\dots,n$. To complete the proof let $\mu_{a_i} = \mu_i$, hence,

$$\sum_{i=0}^{n} |\hat{\mu}_i(\omega)|^2 = \sum_{i=0}^{n} |\hat{\mu}_{a_i}(\omega)|^2 \ge |\hat{\mu}_{a(\omega)}(\omega)|^2 = \prod_{j=1}^{n} \left|\frac{\sin(a(\omega)\omega_j/2)}{a(\omega)\omega_j/2}\right|^2$$

(apply Lemma 9)

$$\ge \left(\frac{8}{5M^4}\right)^{2n} \prod_{j=1}^{n} \frac{1}{\max\left\{\frac{\pi}{M},\ |\omega_j|\right\}^4}. \qquad ∎$$

3.4. Piecewise polynomial approximate identities

Our choice for φ in $\varphi_\tau * f$ is a piecewise polynomial, for φ can then have

i. compact support,

ii. nonnegative values everywhere,

iii. an analytic representation in digital simulations,

iv. a predetermined number of continuous derivatives,

v. a tractable Fourier transform.

Let R be the characteristic function of the unit cube in $\mathbb{R}^n$ centered at the origin. We use the following notation: for any function $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, for $a \in \mathbb{R}^n$, and for $x \in \mathbb{R}^n$,

$$g_s(x) = \frac{1}{s^n}\, g(\frac{x}{s}) \ ,$$

$$g_{[a]}(x) = g(x-a) \ .$$

Our choice for φ is denoted $\varphi_{<k>}$, $k \in \mathbb{N}$,

$$\varphi_{<k>} = (\overbrace{R * R * \cdots * R}^{k+1 \text{ times}})_{1/(k+1)} \ .$$

It has the following readily checked properties:

i. The support of $\varphi_{<k>}$ is the centered unit cube in $\mathbb{R}^n$;

the support of $\varphi_{<k>s}$ is the centered cube of side length s.

$\|\varphi_{<k>s}\|_1 = 1$.

ii. $\varphi_{<k>}(x) \ge 0$, $x \in \mathbb{R}^n$.

iii. $\varphi_{<k>}$ is a piecewise polynomial of degree k.

iv. $\varphi_{<k>}$ has k-1 continuous derivatives.

v. $\hat{\varphi}_{<k>s}(\omega) = \left[\hat{R}\left[\frac{s}{k+1}\,\omega\right]\right]^{k+1}$, $\omega = (\omega_1,\omega_2,\ldots,\omega_n) \in \mathbb{R}^n$,

$$\hat{R}(\omega) = \prod_{j=1}^{n} \frac{\sin(\omega_j/2)}{\omega_j/2} .$$

As usual we have for the first error term ε_1, for $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$,

$$\varepsilon_1 = \left\|\chi_E \left(f - \varphi_{<k>s}*f\right)\right\|_p \xrightarrow[s\to 0]{} 0 \ , \ 1 \le p < \infty,$$

and for $p = \infty$, for x a point of continuity of f,

$$\left|f(x) - \left(\varphi_{<k>s}*f\right)(x)\right| \xrightarrow[s\to 0]{} 0.$$

For any f of interest we can choose a suitable s, but the convergence in not uniform (e.g., f a square wave on $D \subset \mathbb{R}$, $D = \mathrm{supp}\left(\varphi_{<k>s}*\chi_E\right)$, with unit amplitude and period L, then for any fixed s, $p \ne \infty$, ε_1 approaches $\left(\|\chi_E\|_1\right)^{1/p}$ as L approaches 0). Consequently, we have no more to say about any upper bound for ε_1.

We note, however, that for a fixed choice of k, the set $\{\varphi_{<k>s}*f : s>0\}$ is a one parameter subset of $L^p(\mathbb{R}^n)$, and each $\varphi_{<k>}$ is a piecewise polynomial with compact support. These properties make it practical to evaluate by simulation the appropriate size of s for the vision task at hand. Such a choice for s determines ε_1 which in turn suggests an upper bound for ε_2, ε_3, and ε_4 . For $\varepsilon_2+\varepsilon_3+\varepsilon_4$ defines the radius of a neighborhood about $\varphi_{<k>s}*f$. The approach will be to make this radius as small as desired for a fixed s, hence for a fixed ε_1 . A conservative guide would be that the radius should be small compared to ε_1 . Beyond these remarks any additional significance for the size

of the neighborhood depends on additional problem structure such as that discussed in the Introduction. Our interest hereafter is solely how to achieve an error bound radius of a predetermined size.

3.5 INTERPOLATION IN $L^p(E)$

With the choice of $\varphi_{<k>s}$ above we turn to the error ε_2. We shall use frequently the facts that for g and h functions on $\mathbb{R}^n$ such that $g*h$ is defined, for $s>0$ and for $a \in \mathbb{R}^n$,

$$(g*h)_s = g_s * h_s ,$$

$$(g*h)_{[a]} = g*h_{[a]} ,$$

$$\|g_s\|_1 = \|g\|_1 \text{ for } g \in L^1(\mathbb{R}^n), \text{ and}$$

$$\|g*h\|_p \le \|g\|_q \|h\|_r \text{ for } \frac{1}{p} = \frac{1}{q} + \frac{1}{r} - 1,\ 1 \le p,q,r \le \infty,\ 0 \equiv 1/\infty$$

(Young's inequality).

We shall also need

10. LEMMA. For $k \ge 1$, for $y = (y_1, y_2, \ldots, y_n)$, and for $1 \le p \le \infty$,

$$\left\| \left(\varphi_{<k>s}\right)_{[y]} - \varphi_{<k>s} \right\|_p$$
$$\le \left(\frac{k+1}{s}\right)^{n(1-\frac{1}{p})} \min\left\{ 2\frac{k+1}{s}\sum_{i=1}^{n} |y_i| \ ,\ \left(2\frac{k+1}{s}\sum_{i=1}^{n} |y_i|\right)^{1/p} ,\ 2^{1/p} \right\}$$

with the convention $0 = 1/\infty$.

Proof. Define $\mathcal{R} = (\overbrace{R * R * \cdots * R}^{k-1 \text{ times}})_{s/(k+1)}$. To establish the first term in the minimum use

$$\left\| \left(\varphi_{<k>s}\right)_{[y]} - \varphi_{<k>s} \right\|_p = \left\| \mathcal{R} * R_{s/(k+1)} * \left[\left(R_{s/(k+1)}\right)_{[y]} - R_{s/(k+1)}\right] \right\|_p$$
$$\le \|\mathcal{R}\|_1 \ \|R_{s/(k+1)}\|_p \ \left\| \left(R_{s/(k+1)}\right)_{[y]} - R_{s/(k+1)} \right\|_1 ,$$

with the obvious modification if $k = 1$. We have, with $0=1/\infty$,

$$\| \mathcal{R} \|_1 = 1 \qquad \text{and} \qquad \| R_{s/(k+1)} \|_p = \left(\frac{k+1}{s}\right)^{n-\frac{n}{p}},$$

whereas $\| (R_{s/(k+1)})_{[y]} - R_{s/(k+1)} \|_1$ is $\left(\frac{k+1}{s}\right)^n$ times the Lebesgue measure of

$$S = \left(\left[y_1 - \frac{s/2}{k+1}, \, y_1 + \frac{s/2}{k+1}\right] \times \cdots \times \left[y_n - \frac{s/2}{k+1}, \, y_n + \frac{s/2}{k+1}\right] \right) \Delta$$
$$\left(\left[-\frac{s/2}{k+1}, \, \frac{s/2}{k+1}\right] \times \cdots \times \left[-\frac{s/2}{k+1}, \, \frac{s/2}{k+1}\right] \right),$$

where for sets A and B, $A \,\Delta\, B = (A - B) \cup (B - A)$. Let

$$I = \left[-\frac{s/2}{k+1}, \, \frac{s/2}{k+1}\right], \quad I_{y_i} = \left[y_i - \frac{s/2}{k+1}, \, y_i + \frac{s/2}{k+1}\right].$$

Then

$$S \subset \left(\bigcup_{i=1}^{n} I_{y_1} \times \cdots \times I_{y_{i-1}} \times (I_{y_i} - I) \times I_{y_{i+1}} \times \cdots \times I_{y_n} \right) \cup$$
$$\left(\bigcup_{i=1}^{n} I \times \cdots \times I \; (I - I_{y_i}) \times I \times \cdots \times I \right),$$

so that, with $\|S\|$ denoting the measure of S ,

$$\|S\| \le 2 \sum_{i=1}^{n} |y_i| \left(\frac{s}{k+1}\right)^{n-1}.$$

Hence,

$$\| (R_{s/(k+1)})_{[y]} - R_{s/(k+1)} \|_1 \le 2\frac{k+1}{s} \sum_{i=1}^{n} |y_i| ,$$

which completes the proof of the first term in the minimum.

To establish the second and third terms in the minimum use

$$\| (\varphi_{<k>s})_{[y]} - \varphi_{<k>s} \|_p$$
$$\le \|\mathcal{R}\|_1 \; \|R_{s/(k+1)}\|_1 \; \| (R_{s/(k+1)})_{[y]} - R_{s/(k+1)} \|_p .$$

For the case $p<\infty$ the second term follows from

$$\left\| \left(R_{s/(k+1)}\right)_{[y]} - R_{s/(k+1)} \right\|_p \le \left(\frac{k+1}{s}\right)^n \|S\|^{1/p} ,$$

while the third term follows from $\|S\| \le 2\left(\frac{s}{k+1}\right)^n$. For the case $p=\infty$ it suffices for both the second and third terms to note that $\varphi_{<k>s}$ is nonnegative, hence

$$\left\| \left(\varphi_{<k>s}\right)_{[y]} - \varphi_{<k>s} \right\|_\infty \le \left\|\varphi_{<k>s}\right\|_\infty \le \|\mathcal{R}\|_1 \left\|R_{s/(k+1)}\right\|_1 \left\|R_{s/(k+1)}\right\|_\infty \le \left(\frac{k+1}{s}\right)^n .$$

■

The following lemma indicates that we have many choices for an interpolating function.

11. LEMMA. Let $g \in L^1(\mathbb{R}^n)$, $g \ge 0$, and $\operatorname{supp} g \subset B(0,r)$, the ball of radius r in $\mathbb{R}^n$. Let $\tau > 0$, $N \in \mathbb{N}$, and $\bar{N} = (2N+1)^n$. Let $\{e_1, e_2, \ldots, e_n\}$ be an orthonormal basis for $\mathbb{R}^n$, and let $\{x_j\}_{j=1}^{\bar{N}}$ denote the set of points in $\mathbb{R}^n$

$$\left\{ \sum_{i=1}^n p_i \tau e_i \;:\; p_i \in \mathbb{Z},\ |p_i| \le N \right\} .$$

Let $R_{\{\tau\}}(x) = R(\frac{1}{\tau}\, x)$, $x \in \mathbb{R}^n$. Define $\psi = g * R_{\{\tau\}}$ and $\psi_j = \psi_{[x_j]}$. For $E \subset \mathbb{R}^n$ such that $\operatorname{supp}\left(\chi_E * \chi_{B(0,r)}\right) \subset \overbrace{[-N\tau, N\tau] \times \ldots \times [-N\tau, N\tau]}^{n \text{ times}} = D$, then

$$\|g\|_1\, \chi_E = \chi_E \sum_{j=1}^{\bar{N}} \psi_j .$$

Proof. From the definitions

$$\chi_E(x) \sum_{j=1}^{\bar{N}} \psi_j(x) = \chi_E(x) \int_{B(x,r)} g(x-t) \sum_{j=1}^{\bar{N}} \left(R_{\{\tau\}}\right)_{[x_j]}(t)\, dt .$$

It suffices to check that for $x \in E$ then $B(x,r) \subset D$, and

$$\sum_{j=1}^{\bar{N}} \left(R_{\{\tau\}}\right)_{[x_j]} = 1 \text{ on } D .$$

■

Since $R_{\{\tau\}} = \tau^n R_\tau$ by definition,

12. COROLLARY. For $\operatorname{supp}\left(\chi_E * \chi_{B(0,\sqrt{n}\ell\tau)}\right) \subset D$ and for $\psi = \tau^n(\overbrace{R_\tau * \cdots * R_\tau}^{\ell \text{ times}})$, then $\sum_{j=1}^{\bar{N}} \psi_j = 1$ on E.

With this we can now establish an upper bound for ε_2. We choose $\psi = \tau^n(\overbrace{R_\tau * \cdots * R_\tau}^{\ell \text{ times}})$. Let $\{x_j\}$ and ψ_j be as in Lemma 11, and let N be sufficiently large to satisfy the condition in Lemma 11.

13. THEOREM. Let $\varphi = \varphi_{\langle k\rangle s}$, $1 \le q, q' \le \infty$, $1/q + 1/q' = 1$ $(0=1/\infty)$, $f \in L^q(\mathbb{R}^n)$, and let $h = \ell\tau$. For $\varepsilon_2 = \left\| \chi_E\left(\varphi * f - \sum_{j=1}^{\bar{N}} (\varphi * f)(x_j)\ \psi_j\right)\right\|_p$, $1 \le p \le \infty$, then

$$\varepsilon_2 \le \|f\|_q \left(\frac{k+1}{s}\right)^{n/q} \|\chi_E\|_p \min\left\{ \frac{k+1}{s} nh\ ,\ \left(\frac{k+1}{s} nh\right)^{1/q'},\ 2^{1/q'} \right\} .$$

Proof. We have

$$\Big|\chi_E(x)\sum_{j=1}^{\bar N}\int\big(\varphi(x-t)-\varphi(x_j-t)\big)\,f(t)\,dt\,\psi_j(x)\Big|$$

$$\le \chi_E(x)\sum_{j=1}^{\bar N}\|\varphi_{[x-x_j]}-\varphi\|_{q'}\|f\|_q\,\psi_j(x)$$

$$\le \|f\|_q\left(\frac{k+1}{s}\right)^{n(1-\frac{1}{q'})}\min\left\{\frac{k+1}{s}nh\ ,\ \left(\frac{k+1}{s}nh\right)^{1/q'},\ 2^{1/q'}\right\}\chi_E(x)\sum_{j=1}^{\bar N}\psi_j(x)\ ,$$

where the last inequality follows from Lemma 10 and from $\operatorname{supp}\psi_j \subset \{\|x-x_j\|_\infty \le h/2\}$. ■

We conclude this section with some remarks. First, we have required that $f \in L^\infty(\mathbb{R}^n)$ because only for $q = \infty$ does ε_2 depend on s and h according to h/s. This is the simplest case for applications. As we shall see, we will obtain $\|\chi_E\|_p$ as a factor in the bounds for ε_3 and ε_4 as well.

A further remark is that for h/s sufficiently small the minimum has the value of $\frac{k+1}{s}nh$.

A final observation is that the smallest bound is obtained for the choice of $\psi = R_{\{\tau\}}$, that is, $\ell = 1$.

3.6 Approximate reconstruction

In this section we shall determine an explicit upper bound for the third error ε_3. We use the notation and definitions of the Construction section and we use the specific convolutors $\{\mu_i\}_{i=0}^{n}$ of the Lower Bound section. This bound requires more work than any of the others. The first task is to determine the values of k in $\varphi_{<k>s}$ for which $\hat{h}_i = (\varphi_{<k>s})^{\wedge} D_i$ is in $L^2(\mathbb{R}^n)$. To use effectively the lower bound $C(\omega)$ we shall need the following lemmas.

14. Lemma. For a, b, p, q, and x all nonnegative real numbers, for $p-q \geq 0$, and for $b \neq 0$,

$$\frac{(\max\{a,x\})^p}{(\max\{b,x\})^q} \leq \max\left\{\frac{a^p}{b^q}\ ,\ x^{p-q}\right\} .$$

Proof. It suffices to show that the left side of the inequality is bounded by some member of the set on the right hand side for each of the cases: $x \leq a,b$; $a \leq x \leq b$; $b \leq x \leq a$; $a,b \leq x$. ∎

15. Lemma. For a, b, p, q, and x all nonnegative real numbers, for $b \geq a$, $p-q \leq 0$, and for $x \neq 0$, $b \neq 0$,

$$\frac{(\max\{a,x\})^p}{(\max\{b,x\})^q} \leq \min\left\{\frac{(\max\{a,x\})^p}{b^q}\ ,\ x^{p-q}\right\} \leq \min\{b^{p-q}\ ,\ x^{p-q}\} .$$

Proof. For the first inequality it suffices to show that the left hand side is bounded by each term in the set on the right hand side for each of the cases: $x \le a, b$; $a \le x \le b$; $b \le x$. For the second inequality, since $b \ge a$, it suffices to check the cases $x \ge b$ and $x \le b$. ∎

We can now prove

16. **PROPOSITION.** For $(k-2)p > 1$, then $\hat{h}_i = \left(\varphi_{<k>s}\right)^\wedge D_i \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$. In particular, $k = 3$ is sufficient for $\hat{h}_i \in L^2(\mathbb{R}^n)$.

Proof. It is straighforward that for $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n$ then

$$\left(\varphi_{<k>s}\right)^\wedge(\omega) = \prod_{j=1}^{n} \left[\frac{\sin\left[\dfrac{\omega_j s}{2(k+1)}\right]}{\dfrac{\omega_j s}{2(k+1)}} \right]^{k+1} .$$

Since $|\sin(x)| \le \min\{|x|, 1\}$,

$$\left(\varphi_{<k>s}\right)^\wedge(\omega) \le \prod_{j=1}^{n} \min\left\{1 \ , \ 2\frac{k+1}{s} \frac{1}{|\omega_j|}\right\}^{k+1}$$

$$= \left[2\frac{k+1}{s}\right]^{n(k+1)} \prod_{j=1}^{n} \frac{1}{\max\left\{2\dfrac{k+1}{s} \ , \ |\omega_j|\right\}^{k+1}} .$$

From the definitions, $\mu_i = \varphi_{<0>a_i}$. With this and the theorem from the Lower Bound section,

$$|\hat{h}_i(\omega)| = \left| \left(\varphi_{<k>s}\right)^\wedge(\omega) \frac{\overline{\hat{\mu}}_i(\omega)}{\sum_{t=0}^{n} |\hat{\mu}_t(\omega)|^2} \right|$$

$$\le \left(\frac{5M^4}{8}\right)^{2n} \left(2\frac{k+1}{s}\right)^{n(k+1)} \left(\frac{2}{a_i}\right)^n \prod_{j=1}^{n} \frac{\max\left\{\frac{\pi}{M}\ ,\ |\omega_j|\right\}^4}{\max\left\{\frac{2}{a_i}\ ,\ |\omega_j|\right\} \max\left\{2\frac{k+1}{s}\ ,\ |\omega_j|\right\}^{k+1}} .$$

Note that a_i may be replaced by $m = \min\limits_{a_i \in \mathcal{A}} \{a_i\}$. With this and Lemma 14,

$$|\hat{h}_i(\omega)| \le \left(\frac{5M^4}{8}\right)^{2n} \left(\frac{2}{m}\right)^n \left(2\frac{k+1}{s}\right)^{n(k+1)} \prod_{j=1}^{n} \frac{\max\left\{\frac{m}{2}\left(\frac{\pi}{M}\right)^4,\ |\omega_j|^3\right\}}{\max\left\{2\frac{k+1}{s}\ ,\ |\omega_j|\right\}^{k+1}} .$$

Note that $2\frac{k+1}{s}$ may be replaced by $K = \max\left\{2\frac{k+1}{s}\ ,\ \left[\frac{m}{2}\left(\frac{\pi}{M}\right)^4\right]^{1/3}\right\}$. Then by Lemma 15

$$|\hat{h}_i(\omega)| \le \left(\frac{5M^4}{8}\right)^{2n} \left(\frac{2}{m}\right)^n K^{n(k+1)} \prod_{j=1}^{n} \min\left\{\frac{\max\left\{\frac{m}{2}\left(\frac{\pi}{M}\right)^4,\ |\omega_j|^3\right\}}{K^{k+1}}\ ,|\omega_j|^{2-k}\right\}$$

(The case $2\frac{k+1}{s} = K$ is typically the case of interest here.) Hence, $\hat{h}_i \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, whenever $(k-2)p > 1$. ∎

The next result is a well known tool. Our notation for some standard items is: ∂U for the boundary of the set U; i for the imaginary element in $\mathbb{C}$; $\omega \cdot x$ for the usual scalar product of ω and x in $\mathbb{R}^n$; $d\omega$ for the standard volume form for $\mathbb{R}^n$ represented by $d\omega_1 \wedge d\omega_2 \wedge \cdots \wedge d\omega_n$ in the coordinates $(\omega_1, \omega_2, \ldots, \omega_n)$, where $\wedge$ is the wedge product of differential forms; and $\ldots \wedge \widehat{d\omega_j} \wedge \ldots$ for the deletion of the factor $d\omega_j$ in a wedge product.

17. LEMMA. Let $g:\mathbb{R}^n \longrightarrow \mathbb{R}$, $g \in C^k(\mathbb{R}^n)$ (i.e., g has k continuous derivatives). Let $\partial_j g$ denote the partial derivative $\frac{\partial g}{\partial \omega_j}$. Let U be an open set in $\mathbb{R}^n$ with compact closure $\bar{U}$. Let $\bar{U}$ have a triangulation consisting of differentialble singular n-simplexes in $\mathbb{R}^n$. Then

$$\int_U (\partial_{j_1}\partial_{j_2}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\, d\omega$$

$$=\sum_{s=2}^{k-1}\left[\prod_{r=1}^{s-1}(-ix_{j_r})\right](-1)^{j_k+1}\int_{\partial U}(\partial_{j_{s+1}}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\, d\omega_1\wedge\cdots\wedge\widehat{d\omega_{j_s}}\wedge\cdots\wedge d\omega_n$$

$$+\ (-1)^{j_1+1}\int_{\partial U}(\partial_{j_2}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\, d\omega_1\wedge\cdots\wedge\widehat{d\omega_{j_1}}\wedge\cdots\wedge d\omega_n$$

$$+\left[\prod_{r=1}^{k-1}(-ix_{j_r})\right](-1)^{j_k+1}\int_{\partial U} g\ e^{i\omega\cdot x}\, d\omega_1\wedge\cdots\wedge\widehat{d\omega_{j_k}}\wedge\cdots\wedge d\omega_n$$

$$+\left[\prod_{r=1}^{k}(-ix_{j_r})\right]\int_U g\ e^{i\omega\cdot x}\, d\omega\ .$$

Proof. Stokes theorem and induction on k.

18. COROLLARY. Let $\partial_{j_1}\partial_{j_2}\cdots\partial_{j_s} g = 0$ on ∂U for $0 \le s \le k-2$ and for any indices $j_1, j_2,\ldots,j_s$. Then, for $\|x\|$ any norm of $x \in \mathbb{R}^n$,

$$\int_U g\ e^{i\omega\cdot x}\, d\omega = \mathcal{O}\left(\|x\|^{-k}\right) \text{ as } \|x\| \longrightarrow \infty\ .$$

Proof. From the lemma

$$\int_U (\partial_{j_1}\partial_{j_2}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\,d\omega$$

$$- (-1)^{j_1+1}\int_{\partial U}(\partial_{j_2}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\,d\omega_1\wedge\cdots\wedge\widehat{d\omega_{j_1}}\wedge\cdots\wedge d\omega_n$$

$$= \left(\prod_{r=1}^{k}(-ix_{j_r})\right)\int_U g\, e^{i\omega\cdot x}\,d\omega\ .$$

Consequently, by letting $j_1 = 1,2,\ldots,n$ and taking the sum,

$$\int_U \sum_{t=1}^{n}(x_t\partial_t\partial_{j_2}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\,d\omega - \int_{\partial U}(\partial_{j_2}\partial_{j_3}\cdots\partial_{j_k} g)e^{i\omega\cdot x}\,(x\rfloor d\omega)$$

$$= \left(\sum_{t=1}^{n}-ix_t^{\,2}\right)\left(\prod_{r=2}^{k}(-ix_{j_r})\right)\int_U g\, e^{i\omega\cdot x}\,d\omega\ ,$$

where $x\rfloor d\omega = \sum_{j=1}^{n} x_j\,d\omega_1\wedge\cdots\wedge\widehat{d\omega_j}\wedge\cdots\wedge d\omega_n$.

Repeat this for the remaining indices $j_2,\ j_3,\ldots,j_k$ and normalize by dividing by $|x|^k$, $|x|$ the Euclidean norm of x. Let $\partial_{x/|x|} = \sum_{j=1}^{n}\frac{x_j}{|x|}\,\partial_j$.

Thus

$$\int_U(\partial^{k}_{x/|x|}g)e^{i\omega\cdot x}\,d\omega - \int_{\partial U}(\partial^{k-1}_{x/|x|}g)e^{i\omega\cdot x}\,\frac{x}{|x|}\rfloor d\omega)$$

$$= (-i)^k|x|^k\int_U g\, e^{i\omega\cdot x}\,d\omega\ . \qquad \blacksquare$$

We now begin the comparison of $\sum_j(\varphi*f)(x_j)\psi_j$ with an approximation $\sum_j G(x_j)\psi_j$. We consider first $(\varphi*f)(0)$. Some required notation follows. Let $\mathcal{S} = \operatorname{supp}\varphi_{<k>s}$, $\mathcal{M} \supset \bigcup_{i=0}^{n}\operatorname{supp}\mu_i$, let $\chi_{\mathcal{S}}$ and $\chi_{\mathcal{M}}$ deonote the

characteristic functions of $\mathcal{S}$ and $\mathcal{M}$ respectively, and define

$$\mathcal{S} + \mathcal{M} = \operatorname{supp} \chi_{\mathcal{S}} * \chi_{\mathcal{M}} ,$$

$$\mathcal{S} + p\mathcal{M} = \operatorname{supp} \chi_{\mathcal{S}} * \overbrace{\left(\chi_{\mathcal{M}} * \cdots * \chi_{\mathcal{M}}\right)}^{p \text{ times}} .$$

We have $-\mathcal{S} = \mathcal{S}$ and we shall require $-\mathcal{M} = \mathcal{M}$. As usual, we abbreviate $\varphi_{<k>s}$ by φ. Recall that we have the relation

$$\sum_{i=0}^{n} h_i * \mu_i * f = \varphi * f .$$

19. LEMMA.

$$\left(\varphi*f\right)(0) = \left[\sum_{i=0}^{n} h_i * \left((\mu_i * f)\chi_{\mathcal{S}+\mathcal{M}}\right) \right](0)$$
$$+ \left[\sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{S}+2\mathcal{M}}(1-\chi_{\mathcal{S}})\right)\right]\chi_{\mathcal{S}+3\mathcal{M}}(1-\chi_{\mathcal{S}+\mathcal{M}})\right]\right](0) .$$

Proof.

$$\sum_{i=0}^{n} h_i * \left((\mu_i * f)\chi_{\mathcal{S}+\mathcal{M}}\right) = \sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{S}+2\mathcal{M}}(\chi_{\mathcal{S}}+1-\chi_{\mathcal{S}})\right)\right]\chi_{\mathcal{S}+\mathcal{M}}\right]$$
$$= \sum_{i=0}^{n} h_i * \left[\mu_i * \left(f\chi_{\mathcal{S}}\right)\right] + \sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{S}+2\mathcal{M}}(1-\chi_{\mathcal{S}})\right)\right]\chi_{\mathcal{S}+\mathcal{M}}\right] .$$

The first term in this sum evaluated at 0 is

$$\left[\varphi * \left(f\chi_{\mathcal{S}}\right)\right](0) = \varphi * f(0) .$$

The second term evaluated at 0, after adding and subtracting

$$\left[\sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{S}+2\mathcal{M}}(1-\chi_{\mathcal{S}})\right)\right]\chi_{\mathcal{S}+3\mathcal{M}}(1-\chi_{\mathcal{S}+\mathcal{M}})\right]\right](0),$$

is $\left[\sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right]\chi_{\mathcal{G}+3\mathcal{M}}\right]\right](0)$

$$- \left[\sum_{i=0}^{n} h_i * \left[\left[\mu_i * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right]\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}})\right]\right](0),$$

and the first term in this sum is

$$\left[\sum_{i=0}^{n} h_i * \left[\mu_i * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right]\right](0) = \left[\varphi * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right](0) = 0 .$$

■

Now we decompose this expression for $(\varphi*f)(0)$,

$$(\varphi*f)(0) = \left[\sum_{i=0}^{n} (\hat{h}_i\chi_\lambda)^\vee * \left((\mu_i*f)\chi_{\mathcal{G}+\mathcal{M}}\right)\right](0)$$

$$+ \left[\sum_{i=0}^{n} (\hat{h}_i\chi_\lambda)^\vee * \left[\left[\mu_i * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right]\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}})\right]\right](0)$$

$$+ \left[\sum_{i=0}^{n} (\hat{h}_i(1-\chi_\lambda))^\vee * \left((\mu_i*f)\chi_{\mathcal{G}+\mathcal{M}}\right)\right](0)$$

$$+ \left[\sum_{i=0}^{n} (\hat{h}_i(1-\chi_\lambda))^\vee * \left[\left[\mu_i * \left(f\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\right)\right]\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}})\right]\right](0).$$

This same proceedure can be carried out at x_j.

20. COROLLARY.

$$(\varphi*f)(x_j) = \Big[\sum_{i=0}^{n} (\hat{h}_i\chi_\lambda)^\vee * \big((\mu_i*f)(\chi_{\mathcal{G}+\mathcal{M}})_{[x_j]}\big)\Big](x_j)$$

$$+\Big[\sum_{i=0}^{n} (\hat{h}_i\chi_\lambda)^\vee * \Big[\Big[\mu_i* \big(f(\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}}))_{[x_j]}\big)\Big]\Big[\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}})\Big]_{[x_j]}\Big]\Big](x_j)$$

$$+\Big[\sum_{i=0}^{n} (\hat{h}_i(1-\chi_\lambda))^\vee * \big((\mu_i*f)(\chi_{\mathcal{G}+\mathcal{M}})_{[x_j]}\big)\Big](x_j)$$

$$+\Big[\sum_{i=0}^{n} (\hat{h}_i(1-\chi_\lambda))^\vee * \Big[\Big[\mu_i* \big(f(\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}}))_{[x_j]}\big)\Big]\Big[\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}})\Big]_{[x_j]}\Big]\Big](x_j).$$

Proof. Apply $(a*b)(x_j) = (a*b_{[-x_j]})(0)$, e.g.

$$(\varphi*f)(x_j) = (\varphi*f_{[-x_j]})(0) \text{ and}$$

$$(a*(f_{[-x_j]}b))(0) = (a*(fb_{[x_j]}))(x_j) .$$ ∎

The convolutions above can be replaced by the scalar product. Define for a, $b:\mathbb{R}^n \longrightarrow \mathbb{R}$, $\langle a,b\rangle = \int_{\mathbb{R}^n} ab$. Whenever $b(t) = b(-t)$, then $(a*b)(x_j) = \langle a, b_{[x_j]}\rangle = \langle a_{[-x_j]}, b\rangle$. Therefore,

21. COROLLARY.

$$(\varphi*f)(x_j) = \sum_{i=0}^{n} \Big\langle \Big[(\hat{h}_i\chi_\lambda)^\vee\Big]_{[x_j]} , (\mu_i*f)(\chi_{\mathcal{G}+\mathcal{M}})_{[x_j]} \Big\rangle$$

$$+\sum_{i=0}^{n} \Big\langle (\hat{h}_i\chi_\lambda)^\vee , \Big[\mu_i* \big(f_{[-x_j]}\chi_{\mathcal{G}+2\mathcal{M}}(1-\chi_{\mathcal{G}})\big)\Big]\chi_{\mathcal{G}+3\mathcal{M}}(1-\chi_{\mathcal{G}+\mathcal{M}}) \Big\rangle$$

$$+\sum_{i=0}^{n} \Big\langle (\hat{h}_i(1-\chi_\lambda))^\vee ,$$

$$\left(\mu_i * f_{[-x_j]}\right)\chi_{\mathcal{S}+\mathcal{M}} + \left[\mu_i * \left(f_{[-x_j]}\chi_{\mathcal{S}+2\mathcal{M}}(1-\chi_{\mathcal{S}})\right)\right]\chi_{\mathcal{S}+3\mathcal{M}}(1-\chi_{\mathcal{S}+\mathcal{M}}) >.$$

Let $\eta_1(x_j)$ and $\eta_2(x_j)$ be the second and third terms respectively of the right hand side above. Then

$$\varepsilon_3 =$$

$$\left\|\chi_E\left[\sum_{j\in J}(\varphi * f)(x_j)\psi_j - \sum_{j\in J}\sum_{i=0}^{n} < \left((\hat{h}_i\chi_\lambda)^\vee\right)_{[x_j]}, \; (\chi_{\mathcal{S}+\mathcal{M}})_{[x_j]}(\mu_i * f) > \psi_j\right]\right\|_p$$

$$\le \|\chi_E \sum_{j\in J}\eta_1(x_j)\psi_j\|_p + \|\chi_E\sum_{j\in J}\eta_2(x_j)\psi_j\|_p$$

$$\le \max\left\{|\eta_1(x_j)|\right\}\|\chi_E\|_p + \max\left\{|\eta_2(x_j)|\right\}\|\chi_E\|_p .$$

We turn to determining bounds for $\eta_1(x_j)$ and $\eta_2(x_j)$. We address $\eta_2(x_j)$ first, this case being easier.

First, $\eta_2(x_j)$ is the sum of two terms, each of the form $\sum_{i=0}^{n} < \left(\hat{h}_i(1-\chi_\lambda)\right)^\vee , \; \chi_S\left(\mu_i * \left(f_{[-x_j]}\chi_T\right)\right) >$, which is bounded by

$$\sum_{i=0}^{n}\|\left(\hat{h}_i(1-\chi_\lambda)\right)^\vee\|_2 \; \|\chi_S\left(\mu_i * \left(f_{[-x_j]}\chi_T\right)\right)\|_2$$

$$\le \sum_{i=0}^{n}(2\pi)^{-n/2}\;\|\hat{h}_i(1-\chi_\lambda)\|_2 \; \|\mu_i * \left(f_{[x_j]}\chi_T\right)\|_2$$

$$\le \sum_{i=0}^{n}(2\pi)^{-n/2}\;\|\hat{h}_i(1-\chi_\lambda)\|_2 \; \|\mu_i\|_1\|f\|_2 .$$

It is easy to see that this also bounds $\eta_2(x_j)$. Recall that $\|\mu_i\|_1 = 1$.

We use the proof of Proposition 16 to bound $\|\left(\hat{h}_i(1-\chi_\lambda)\right)^\vee\|_p$, $p>1$. If $s\le M$ and $k\ge 3$ then $2\frac{k+1}{s} \ge \left[\frac{m}{2}\left(\frac{\pi}{M}\right)^4\right]^{1/3}$, and this is the case we shall assume. Let

$$a = \left[\frac{m}{2}\left(\frac{\pi}{M}\right)^4\right]^{1/3}, \qquad b = 2\frac{k+1}{s}$$

so that by the proof of Proposition 16, with C a constant,

$$|\hat{h}_i(\omega)| \le C \prod_{j=1}^{n} \min\left\{\frac{\max\left\{a^3, |\omega_j|^3\right\}}{b^{k+1}}, |\omega_j|^{2-k}\right\}.$$

By definition, $1-\chi_\lambda$ is the characteristic function of the set $\{\|\omega\|_\infty > \lambda\}$, where $\|\omega\|_\infty = \max\{|\omega_j|, j=1,2,\ldots n\}$. Note that the zero set of $(\varphi_{<k>s})^\wedge$ contains $\{\|\omega\|_\infty = \ell\pi\, 2\frac{k+1}{s}, \ell\in\mathbb{N}\}$. Consequently, because of Lemma 17 and its corollary, we shall later choose $\lambda = \ell\pi\, 2\frac{k+1}{s} = \ell\pi b$, and this is convenient here also. Finally observe that

$$\{\|\omega\|_\infty > \lambda\} = \bigcup_{i=1}^{n} \{\|\omega\|_\infty > \lambda\} \cap \{|\omega_i| = \|\omega\|_\infty\}.$$

We outline the integration.

$$C^{-p}\|\hat{h}_i(1-\chi_\lambda)\|_p^p \le \sum_{i=1}^{n} \int\limits_{\{\|\omega\|_\infty>\lambda\}\cap\{|\omega_i|=\|\omega\|_\infty\}} \prod_{j=1}^{n} \min\{\cdots\}^p \le n \int\limits_{\{\|\omega\|_\infty>\lambda\}\cap\{|\omega_n|=\|\omega\|_\infty\}} \prod_{j=1}^{n} \min\{\cdots\}^p$$

$$= n2^n \int_\lambda^\infty y^{p(2-k)}\left[\int_0^a \left(\frac{a^3}{b^{k+1}}\right)^p dx + \int_a^b \left(\frac{x^3}{b^{k+1}}\right)^p dx + \int_b^y x^{p(2-k)} dx\right]^{n-1}$$

(use $a\le b$ and the inequality $|v^n - (v+u)^n| \le n|u|\left(|v|+|u|\right)^{n-1}$)

$$\le n2^n \left(b^{1+p(2-k)}\right)^n \left(1 + \frac{1+(\ell\pi)^{1+p(2-k)}}{p(k-2)-1}\right)^{n-1} \frac{(\ell\pi)^{1+p(2-k)}}{p(k-2)-1}.$$

With this and with $p = 2$,

$$\max_j\{|\eta_2(x_j)|\}$$
$$\leq (n+1)\|f\|_2\left(\frac{5M^4}{8}\right)^{2n}\left(\frac{2}{m}\right)^n\left(2\frac{k+1}{s}\right)^{\frac{7n}{2}}\left(\frac{n}{\pi^n}\right)^{\frac{1}{2}}$$
$$\times\left[\left(1+\frac{1+(\ell\pi)^{1+2(2-k)}}{2(k-2)-1}\right)^{n-1}\frac{(\ell\pi)^{1+2(2-k)}}{2(k-2)-1}\right]^{\frac{1}{2}}.$$

We now turn to $\eta_1(x_j)$. Whereas we used λ to control the size of $\eta_2(x_j)$, we shall depend on M to control the size of

$$\eta_1(x_j) = \sum_{i=0}^{n} <\left(\hat{h}_i\chi_\lambda\right)^\vee, \left[\mu_i*\left(f_{[-x_j]}\chi_{\varphi+2M}(1-\chi_\varphi)\right)\right]\chi_{\varphi+3M}(1-\chi_{\varphi+M})>.$$

How this is done is indicated in Lemma 22 below. First some notation. As previously noted, we use χ_λ to denote the characteristic function of $\{\|\omega\|_\infty\leq\lambda\}$ and we choose $\lambda = \ell\pi 2\frac{k+1}{s}$, where ℓ is some positive integer. For simplicity, let $\Lambda = \{\|\omega\|_\infty\leq\lambda\}$. Then

$$\Lambda = \bigcap_{i=1}^{n}\{|\omega_j| \leq \lambda\},$$
$$\partial\Lambda \subset \bigcup_{i=1}^{n}\{|\omega_j| = \lambda\}.$$

A second item of notation is the multi-index $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_n) \in \mathbb{N}^n$. Define $|\alpha| = \sum_{j=1}^{n}\alpha_j$, $\alpha! = \alpha_1!\alpha_2!\cdots\alpha_n!$, and $\partial^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2}\cdots\partial_n^{\alpha_n}$.

22. LEMMA. If $|\alpha| \leq k$, then $\partial^\alpha\hat{h}_i(\omega) = 0$ for $\omega \in \partial\Lambda$. Consequently, for all $r\leq k+1$, $x \in \mathbb{R}^n$, with $|x|$ the Euclidean norm of x,

$$\left|\left(\hat{h}_i\chi_\lambda\right)^\vee(x)\right| \leq \frac{1}{(2\pi)^n}\frac{1}{|x|^r}\left\|\partial_{x/|x|}^{\,r}\hat{h}_i\right\|_{1,\Lambda}.$$

Proof. The first statement follows from the property v of $\varphi_{<k>s}$ in Section 4, from the product formulas for derivatives, and from the definition of h_i. The second statement follow from the proof of Corollary 18. ∎

Several lemmas will be required to bound $\|\partial^r_{x/|x|}\hat{h}_i\|_{1,\Lambda}$. Since $\hat{h}_i(\omega) = (\varphi_{<k>s})^\wedge(\omega)\ D_i(w)$, it suffices to bound $\partial^\alpha\left[(\varphi_{<k>s})^\wedge\right](\omega)$ and $\partial^r_{x/|x|}D_i(\omega)$, to apply Leibnitz's rule, and finally to integrate. Recall Leibnitz's rule for $\mathbb{C}$ valued functions f and g on $\mathbb{R}^n$, with α, β, γ multi-indices:

$$\partial^\alpha(fg) = \sum_{\substack{\beta,\gamma \\ \beta+\gamma=\alpha}} \frac{\alpha!}{\beta!\gamma!}(\partial^\beta f)(\partial^\gamma g)\ .$$

(In the particular cases examined here all Fourier transforms have values only on the real axis.)

It will be convenient to have a bound for $\partial^\alpha\left[(\varphi_{<k>s})^\wedge\right](\omega)$ which depends only on $|\alpha|$.

23. LEMMA. $\left|\partial^\alpha\left[(\varphi_{<k>s})^\wedge\right](\omega)\right| \le$

$$\left[\frac{s}{2(k+1)}\right]^{|\alpha|}\frac{(k+|\alpha|)!}{k!}\,2^{|\alpha|}\prod_{j=1}^{n}\min\left\{1\ ,\frac{1}{|\omega_j|\frac{s}{2(k+1)}}\right\}^{k+1}.$$

($\min\{1,1/x\}$ is understood to be 1 for $x = 0$.)

Proof. Let $\varphi'_{<k>s}$ denote $\varphi_{<k>s}$ constructed for $\mathbb{R}$. From the definitions, for $\omega = (\omega_1,\omega_2,\dots,\omega_n) \in \mathbb{R}^n$,

$$\left(\varphi_{<k>s}\right)^{\wedge}(\omega) = \prod_{j=1}^{n} \left(\varphi'_{<k>s}\right)^{\wedge}(\omega_j) \ .$$

Claim. For $\upsilon \in \mathbb{R}$ and for $r \in \mathbb{N}$,

$$|\partial^r\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\upsilon)| \le \left(\frac{s}{2}\right)^r \min\left\{ 1 \ , \ \frac{(k+r)!}{k!\,(k+1)^r} \left(1 + \frac{2(k+1)}{|\upsilon|s}\right)^r \left(\frac{2(k+1)}{|\upsilon|s}\right)^{k+1}\right\} .$$

Proof of Claim. The first element in the minimized set is established by

$$|\partial^r\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\upsilon)| \le \int_{[-\frac{s}{2},\frac{s}{2}]} |\varphi'_{<k>s}(t)|\,|t|^r dt \le \left(\frac{s}{2}\right)^r \|\varphi'_{<k>s}\|_1 \ .$$

The second element is established by induction on k. Recall

$$\left(\varphi'_{<k>s}\right)^{\wedge}(\upsilon) = \left[\operatorname{sinc}\left(\frac{\upsilon s}{2(k+1)}\right)\right]^{k+1} , \quad \operatorname{sinc}(\upsilon) = \sin(\upsilon)/\upsilon \ .$$

By Leibnitz's rule

$$|\partial^r\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\upsilon)| \le \sum_{e=0}^{r} \binom{r}{e} \left(\frac{s}{2}\right)^{r-e} \frac{e!}{|\upsilon|^{e+1}}\,\frac{2}{s} \le r!\left(\frac{s}{2}\right)^r \left(1 + \frac{2}{|\upsilon|s}\right)^r \left(\frac{2}{|\upsilon|s}\right),$$

hence the result for $k = 0$. Recall the notation $g_{\{s\}}(\upsilon) = g(\upsilon/s)$ for $s>0$. Consequently $(\partial^r g_{\{s\}})(\upsilon) = (1/s)^r(\partial^r g)(\upsilon/s)$. With this notation

$$\left(\varphi'_{<k>s}\right)^{\wedge} = \left[\left(\varphi'_{<k-1>s}\right)^{\wedge}\right]_{\{\frac{k+1}{k}\}} \cdot \left[\left(\varphi'_{<0>s}\right)^{\wedge}\right]_{\{k+1\}} \ .$$

From Leibnitz's rule, from the result for $k=0$, and from the induction hypothesis

$$|\partial^r\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\upsilon)| \le \left(\frac{s}{2}\right)^r \frac{1}{(k+1)^r}\left(1 + \frac{2(k+1)}{|\upsilon|s}\right)^r \left(\frac{2(k+1)}{|\upsilon|s}\right)^{k+1} \sum_{e=0}^{r} \binom{r}{e} \frac{(e+k-1)!}{(k-1)!}(r-e)! \ ,$$

and the final sum is checked by induction to be $\frac{(k+r)!}{k!}$. This proves the claim.

Observe that $\dfrac{(k+r)!}{k!\,(k+1)^r} \geq 1$ and that for y>0

$$\min\{1\ ,\ (1+y)^r y^{k+1}\} \leq \min\{1\ ,\ 2^r y^{k+1}\} \leq 2^r \min\{1\ ,\ y^{k+1}\}\ .$$

With these observations

$$\left|\partial^r\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\upsilon)\right| \leq$$

$$\left[\frac{s}{2(k+1)}\right]^r \frac{(k+r)!}{k!}\, 2^r \prod_{j=1}^{n} \min\left\{\ 1\ ,\frac{1}{|\upsilon|\frac{s}{2(k+1)}}\right\}^{k+1}.$$

Since $\partial^{\alpha}\left[\left(\varphi_{<k>s}\right)^{\wedge}\right](\omega) = \prod_{j=1}^{n} \partial^{\alpha_j}\left[\left(\varphi'_{<k>s}\right)^{\wedge}\right](\omega_j)$, it remains only to check by induction on n that for a multi-index α

$$\prod_{j=1}^{n} \frac{(k+\alpha_j)!}{k!} \leq \frac{k+|\alpha|)!}{k!}\ .$$

∎

We next bound derivatives of $D_i = \bar{\hat{\mu}}_i \left[\sum_{i=0}^{n} |\hat{\mu}_i|^2\right]^{-1}$. From Leibnitz's rule and since $\hat{\mu}_i = \left(\varphi_{<0>a_i}\right)^{\wedge}$, it suffices to consider derivatives of the second factor. A formula for higher derivatives of compositions of functions will be needed. Let s(e) denote a multi-index with e coordinates $s(e) = (s_1, s_2, \ldots, s_e)$. For a function f of one variable let $f^{(s_i)}$ denote the derivative of order s_i.

24. LEMMA. For $g, f \in C^r(\mathbb{R})$, with $r \geq 1$,

$$(f\circ g)^{(r)} = \sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \begin{bmatrix} s_1+\cdots+s_e-1 \\ s_1-1 \end{bmatrix} \begin{bmatrix} s_2+\cdots+s_e-1 \\ s_2-1 \end{bmatrix} \cdots \begin{bmatrix} s_e-1 \\ s_e-1 \end{bmatrix}$$

$$\times \left[f^{(e)}\circ g\right]\left[g^{(s_1)} g^{(s_2)} \cdots g^{(s_e)}\right] .$$

Proof. The proof is by induction on r. With the convention $0!=1$ the case $r=1$ is clear. Assume the result for $r-1$.

$$(f\circ g)^{(r)} = \left[(f\circ g)^{(1)}\right]^{(r-1)} = \left[\left(f^{(1)}\circ g\right)\cdot g^{(1)}\right]^{(r-1)}$$

(apply Leibnitz's rule and the induction hypothesis)

$$= \sum_{s=1}^{r-1} \begin{bmatrix} r-1 \\ s-1 \end{bmatrix} \sum_{e=1}^{r-s} \sum_{\substack{|s(e)|=r-s \\ s_i \geq 1}} \begin{bmatrix} s_1+\cdots+s_e-1 \\ s_1-1 \end{bmatrix} \begin{bmatrix} s_2+\cdots+s_e-1 \\ s_2-1 \end{bmatrix} \cdots \begin{bmatrix} s_e-1 \\ s_e-1 \end{bmatrix}$$

$$\times \left[f^{(e+1)}\circ g\right]\left[g^{(s_1)} g^{(s_2)} \cdots g^{(s_e)} g^{(s)}\right] + \left(f^{(1)}\circ g\right)\cdot g^{(r)} .$$

Observe

$$\sum_{s=1}^{r-1} \begin{bmatrix} r-1 \\ s-1 \end{bmatrix} \sum_{e=1}^{r-s} \sum_{\substack{|s(e)|=r-s \\ s_i \geq 1}} \left[\cdots\right] = \sum_{e=1}^{r-1} \sum_{s=1}^{r-e} \sum_{\substack{|s(e)|=r-s \\ s_i \geq 1}} \begin{bmatrix} r-1 \\ s-1 \end{bmatrix} \left[\cdots\right]$$

$$= \sum_{e=1}^{r-1} \sum_{\substack{s_0+s_1+s_2+\cdots+s_e=r \\ s_0, s_i \geq 1}} \begin{bmatrix} r-1 \\ s-1 \end{bmatrix} \left[\cdots\right] .$$

The term $\left(f^{(1)}\circ g\right)\cdot g^{(r)}$ corresponds to an additional $e=0$ term in the last formulation of the summation. By renaming these r values for the

index e (add 1) the desired form is obtained. ∎

Some miscellaneous results that will be needed are collected in

25. LEMMA. Let ∂_v be the directional derivative in direction $v \in \mathbb{R}^n$.

i. Lemma 24 holds for $f \in C(\mathbb{R})$, $g \in C(\mathbb{R}^n)$ if ∂^r is replaced by ∂_v^r.

ii.

$$\sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \binom{s_1+\cdots+s_e-1}{s_1-1}\binom{s_2+\cdots+s_e-1}{s_2-1}\cdots\binom{s_e-1}{s_e-1}$$

$$= \sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \binom{r-1}{s_1-1}\binom{r-1-s_1}{s_2-1}\cdots\binom{r-1-s_1-\cdots-s_{e-1}}{s_e-1}$$

$$\leq r^r.$$

iii. For $\varphi \in C^k(\mathbb{R}^n)$ and for $r=|\alpha|\leq k$, if $M(|\alpha|)$ is a bound for $\partial^\alpha\varphi$ which depends only on $|\alpha|$, then, for $|v|$ the Euclidean norm of v,

$$|\partial_v^r\varphi| \leq (\sqrt{n})^r|v|^r M(r) .$$

Proof. For *i*, if $f \in C(\mathbb{R}^n)$, $x, v \in \mathbb{R}^n$, and if $p_v:\mathbb{R} \longrightarrow \mathbb{R}^n$, $p_v(t) = x + vt$, then $\left(\partial_v^r f\right)(x) = \left(\partial^r(f\circ p_v)\right)(0)$.

For *ii*, The first relation uses $s_1+s_2+\cdots+s_e = r$. The inequality follows from

$$\sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \frac{r}{r} \frac{(r-1)!}{(s_1-1)!(r-s_1)!} \frac{(r-s_1-1)!}{(s_2-1!(r-s_1-s_2)!} \frac{(r-s_1-s_2-1)!}{(s_3-1)!(r-s_1-s_2-s_3)!} \cdots$$

$$\times \frac{(r-s_1-s_2-\cdots-s_{e-1}-1)!}{(s_e-1)!0!}$$

$$= \sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \frac{r!}{\left[r(r-s_1)(r-s_1-s_2)\cdots(r-s_1-\cdots-s_{e-1})\right]}$$

$$\times \frac{1}{\left[(s_1-1)!(s_2-1)!\cdots(s_e-1)!\right]}$$

$$= \sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \frac{r!}{\left[r(s_1-1)!\right]\left[(r-s_1)(s_2-1)!\right]\cdots\left[(r-s_1-\cdots-s_{e-1})(s_e-1)!\right]}$$

$$\leq \sum_{e=1}^{r} \sum_{\substack{|s(e)|=r \\ s_i \geq 1}} \frac{r!}{s_1!s_2!\cdots s_e!} \leq \sum_{|\alpha|=r} \frac{r!}{\alpha!} ,$$ with this last summation

in multi-index notation ($\alpha \in \mathbb{N}^r$) , while for $m \in \mathbb{N}$, $s_i \in \mathbb{R}$, $i=1,2,\ldots,r$,

$$\left[\sum_{i=1}^{r} s_i\right]^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} s_1^{\alpha_1} s_2^{\alpha_2} \cdots s_r^{\alpha_r} .$$

For *iii*, use

$$\sum_{i=1}^{n} |v_i| \leq \sqrt{n}|v| , \quad \partial_v \varphi = \sum_{i=1}^{n} v_i \partial_i \varphi ,$$

hence

$$|\partial_{\mathrm{v}}^{r}\varphi| = \Big|\sum_{t_1=1}^{n} v_{t_1}\partial_{t_1}\cdots\sum_{t_r=1}^{n} v_{t_r}\partial_{t_r}\varphi\Big| \le \sum_{t_1=1}^{n}\cdots\sum_{t_r=1}^{n}|v_{t_1}\cdots v_{t_r}|\, M(r)$$
$$\le (\sqrt{n}|\mathrm{v}|)^{r}M(r). \qquad \blacksquare$$

Now we can complete the bound for $|\partial_{\mathrm{v}}^{r}D_i(\omega)|$.

26. LEMMA. For $\mathrm{v} \in \mathbb{R}^n$, $|\mathrm{v}|=1$, for $r\ge1$, and for $\omega = (\omega_1,\omega_2,\ldots,\omega_n) \in \mathbb{R}^n$,

$$|\partial_{\mathrm{v}}^{r}D_i(\omega)| \le \big(M\sqrt{n}(n+1)r\big)^{r}(r+1)!\,\left(\frac{5M^4}{8}\right)^{2n(r+1)}\left(\frac{2}{m}\right)^{n(2r+1)}$$
$$\times\prod_{j=1}^{n}\max\left\{\left(\frac{\pi}{M}\right)^{4(r+1)}\left(\frac{m}{2}\right)^{2r+1},\ |\omega_j|^{2r+3}\right\}.$$

Proof. From Leibnitz's rule and from Lemma 24 with $f(t) = t^{-1}$ and $g = \sum_{i=0}^{n}|\hat{\mu}_i|^2$,

$$\partial_{\mathrm{v}}^{r}D_i = \sum_{p=0}^{r}\binom{r}{p}\partial_{\mathrm{v}}^{r-p}\bar{\hat{\mu}}_i\sum_{e=1}^{p}\sum_{\substack{|s(e)|=p\\ s_i\ge1}}\binom{s_1+\cdots+s_e-1}{s_1-1}\cdots\binom{s_e-1}{s_e-1}\frac{(-1)^e e!}{\left(\sum_{i=0}^{n}|\hat{\mu}_i|^2\right)^{e+1}}$$
$$\times\prod_{q=1}^{e}\sum_{j=0}^{n}\sum_{t_q=0}^{s_q}\binom{s_q}{t_q}\left(\partial_{\mathrm{v}}^{s_q-t_q}\hat{\mu}_i\right)\left(\partial_{\mathrm{v}}^{t_q}\bar{\hat{\mu}}_i\right).$$

From Lemma 23 with $\mu_i = \varphi_{\langle 0\rangle a_i}$, $m = \min\{a_i\}$, $M = \max\{a_i\}$, and from Lemma 25,

$$|\partial_{\mathrm{v}}^{\mathrm{r}} D_{\mathrm{i}}(\omega)| \le \sum_{\mathrm{p}=0}^{r} \binom{r}{\mathrm{p}} (\sqrt{n}M)^{\mathrm{r-p}} (r-\mathrm{p})! \prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1\ ,\ \frac{2}{|\omega_{\mathrm{j}}|m}\right\}$$

$$\times \sum_{\mathrm{e}=1}^{\mathrm{p}} \sum_{\substack{|\mathrm{s(e)}|=\mathrm{p} \\ \mathrm{s_i}\ge 1}} \begin{bmatrix} \mathrm{s_1}+\cdots+\mathrm{s_e}-1 \\ \mathrm{s_1}-1 \end{bmatrix} \cdots \begin{bmatrix} \mathrm{s_e}-1 \\ \mathrm{s_e}-1 \end{bmatrix} \frac{(-1)^{\mathrm{e}}\mathrm{e}!}{\left(\sum_{\mathrm{i}=0}^{\mathrm{n}} |\hat{\mu}_{\mathrm{i}}|^2\right)^{\mathrm{e+1}}}$$

$$\times \prod_{\mathrm{q}=1}^{\mathrm{e}} \sum_{\mathrm{j}=0}^{\mathrm{n}} \sum_{\mathrm{t_q}=0}^{\mathrm{s_q}} \begin{bmatrix} \mathrm{s_q} \\ \mathrm{t_q} \end{bmatrix} (\sqrt{n}M)^{\mathrm{s_q}} (\mathrm{s_q}-\mathrm{t_q})!\,\mathrm{t_q}! \prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1\ ,\ \frac{2}{|\omega_{\mathrm{j}}|m}\right\}^2 .$$

For the last factor,

$$\prod_{\mathrm{q}=1}^{\mathrm{e}} \sum_{\mathrm{j}=0}^{\mathrm{n}} \sum_{\mathrm{t_q}=0}^{\mathrm{s_q}} \left[\cdots\right] \le \prod_{\mathrm{q}=1}^{\mathrm{e}} (n+1)(\mathrm{s_q}+1)!\,(\sqrt{n}M)^{\mathrm{s_q}} \prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1\ ,\ \frac{2}{|\omega_{\mathrm{j}}|m}\right\}^2$$

(using $\prod_{\mathrm{j}=1}^{\mathrm{n}} (\alpha_{\mathrm{j}}+1)! \le (|\alpha|+1)!$ as in the proof of Lemma 23)

$$\le (n+1)^{\mathrm{e}} (\sqrt{n}M)^{\mathrm{p}} \prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1\ ,\ \frac{2}{|\omega_{\mathrm{j}}|m}\right\}^{2\mathrm{e}} (\mathrm{p}+1)!\ .$$

Combining,

$$|\partial_{\mathrm{v}}^{\mathrm{r}} D_{\mathrm{i}}(\omega)| \le r!\,(\sqrt{n}M)^{\mathrm{r}} \frac{\prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1\ ,\ \frac{2}{|\omega_{\mathrm{j}}|m}\right\}}{\sum_{\mathrm{i}=0}^{\mathrm{n}} |\hat{\mu}_{\mathrm{i}}(\omega)|^2} \sum_{\mathrm{p}=0}^{r} (\mathrm{p}+1)$$

$$\times \sum_{\mathrm{e}=1}^{\mathrm{p}} \sum_{\substack{|\mathrm{s(e)}|=\mathrm{p} \\ \mathrm{s_i}\ge 1}} \begin{bmatrix} \mathrm{s_1}+\cdots+\mathrm{s_e}-1 \\ \mathrm{s_1}-1 \end{bmatrix} \cdots \begin{bmatrix} \mathrm{s_e}-1 \\ \mathrm{s_e}-1 \end{bmatrix} \mathrm{e}! \left(\frac{(n+1) \prod_{\mathrm{j}=1}^{\mathrm{n}} \min\left\{1, \frac{2}{|\omega_{\mathrm{j}}|m}\right\}}{\sum_{\mathrm{i}=0}^{\mathrm{n}} |\hat{\mu}_{\mathrm{i}}(\omega)|^2} \right)^{\mathrm{e}} .$$

Since $|\hat{\mu}_i(\omega)|^2 \le \prod_{j=1}^{n} \min\left\{1\ ,\ \frac{2}{|\omega_j|m}\right\}^2$,

$$|\partial_v^r D_i(\omega)| \le r!(\sqrt{n}M)^r \sum_{p=0}^{r} (p+1)!p^p(n+1)^p \frac{\prod_{j=1}^{n} \min\left\{1\ ,\ \frac{2}{|\omega_j|m}\right\}^{2p+1}}{\left[\sum_{i=0}^{n} |\hat{\mu}_i(\omega)|^2\right]^{p+1}}$$

$$\le ((r+1)!)^2(\sqrt{n}M)^r(n+1)^r \frac{\prod_{j=1}^{n} \min\left\{1\ ,\ \frac{2}{|\omega_j|m}\right\}^{2r+1}}{\left[\sum_{i=0}^{n} |\hat{\mu}_i(\omega)|^2\right]^{r+1}} .$$

To complete the proof apply the lower bound theorem, then apply Lemma 1. ■

At last we bound $|\partial_v^r \hat{h}_i(\omega)|$.

27. LEMMA. For $v \in \mathbb{R}^n$, $|v|=1$, $r \ge 1$, and for $\omega = (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n$,

$$|\partial_v^r \hat{h}_i(\omega)| \le A(r,k,n,\mathcal{A}) \left(2\frac{k+1}{s}\right)^{n(k+1)} \frac{\prod_{j=1}^{n} \max\{\ C\ , |\omega_j|\}^{2r+3}}{\prod_{j=1}^{n} \max\left\{2\frac{k+1}{s}\ , |\omega_j|\ \right\}^{k+1}} ,$$

where $C = \max\limits_{e=0,1,\ldots,r} \left\{\left[\left(\frac{\pi}{M}\right)^{4(e+1)} \left(\frac{m}{2}\right)^{2e+1}\right]^{1/(2e+3)}\right\}$, and where

$$A(r,k,n,\mathcal{A}) = ((r+1)!)^2\binom{k+r}{r}\left(\frac{5M^4}{8}\right)^{2n}\left(\frac{2}{m}\right)^n\left[\frac{s}{(k+1)C^{2n}} + \sqrt{n}M(n+1)r\left(\frac{5M^4}{8}\ \frac{2}{m}\right)^{2n}\right]^r$$

with $\mathcal{A}$ the set $\{a_0, a_1, \ldots, a_n\}$ of which M and m are the max and min, respectively.

Proof. Apply Leibnitz's rule, Lemma 26, and Lemma 23, then use the inequalities

$$\max\left\{\left(\frac{\pi}{M}\right)^{4(e+1)}\left(\frac{m}{2}\right)^{2e+1}, |\omega_j|^{2e+3}\right\} \le \max\left\{C, |\omega_j|\right\}^{2e+3}$$

$$C^{2e+3}\max\left\{1, |\omega_j|/C\right\}^{2r+3}, \text{ for } e \le r,$$

and

$$((e+1)!)^2 \frac{(k+r-e)!}{k!} \le (e+1)!(e+1)\frac{(k+r)!}{k!} \le ((r+1)!)^2\binom{k+r}{r},$$

and finally apply the binomial formula. ∎

We return to our original goal, the bound for $|\eta_1(x_j)|$. Lemma 22 and Lemma 27 narrow the choices for r and k. By Hölder's inequality and by Young's inequality, along with $\|\mu_i\|_1 = 1$,

$$|\eta_1(x_j)| \le \sum_{i=0}^{n} \|(\hat{h}_i\chi_\lambda)^\vee(1-\chi_{\mathcal{M}})\|_2\|f\|_2 .$$

We may assume that $\mathcal{M}$ is a centered cube in $\mathbb{R}^n$ with side length β. Let $B_i(r,k,n,\lambda)$ bound $\|\partial^r_{x/|x|}\hat{h}_i\|_{1,\Lambda}$ uniformly in $x \in \mathbb{R}^n-\{0\}$. Then by Lemma 22,

$$\|(\hat{h}_i\chi_\lambda)^\vee(1-\chi_{\mathcal{M}})\|_2 \le \frac{B_i(r,k,n,\lambda)}{(2\pi)^n}\left[\int_{\mathbb{R}^n-\mathcal{M}} |x|^{-2r}dx\right]^{1/2}.$$

Consequently, we choose $2r \ge n+1$. Since $\|x\|_\infty \le |x|$, we integrate over $\{\|x\|_\infty > \beta\}$ in the manner as that outlined in the discussion preceding

Lemma 22:

$$\|(\hat{h}_1\chi_\lambda)^\vee(1-\chi_{\mathcal{M}})\|_2 \le \frac{B_1(r,k,n,\lambda)}{(2\pi)^n}\ \frac{\sqrt{n}}{\sqrt{2r-n}}\ 2^r\ \left(\frac{1}{\beta}\right)^{\frac{2r-n}{2}} .$$

From Lemma 27 $B_1(r,k,n,\lambda)$ can be determined.

28. LEMMA. For v, r, ω, $\mathcal{A}$, C, and A as in Lemma 27, for $k-2r-2\neq 1$, for $\lambda = \ell\pi 2^{\frac{k+1}{s}}$, $\ell \in \mathbb{N}-\{0\}$, $\Lambda = \{\|\omega\|_\infty \le \lambda\}$,

$$\|\partial_v^r\hat{h}_1\|_{1,\Lambda} \le A(r,k,n,\mathcal{A})\ 2^n\ K_1^{2n(r+2)}\left[\min\{1,\frac{\lambda}{K_1}\}^n + u\left[\left(\frac{N+(\ell\pi)^{N+1}}{N+1}\right)^n-1\right]\right]$$

where

$$N = 2r+2-k, \quad u = \begin{cases} 0 \text{ if } \lambda\le K_1 \\ 1 \text{ if } \lambda > K_1 \end{cases}, \quad K_1 = \max\left\{2^{\frac{k+1}{s}},\ C\right\} .$$

Proof. Note that each occurrence of $2^{\frac{k+1}{s}}$ in the first inequality of Lemma 27 may be replaced by K_1. Thus, by Lemmas 14 and 15 combined,

$$|\partial_v^r\hat{h}_1(\omega)| \le A(r,k,n,\mathcal{A})K_1^{n(k+1)}\prod_{j=1}^n \max\{K_1,|\omega_j|\}^{2r+3-k-1} .$$

Integrate, treating separately $\|\omega\|_\infty \le K_1$ and $\|\omega\|_\infty > K_1$. For example, for $\lambda \ge K_1$ and using $\mathbb{R}^n = \bigcup_{i=1}^n \{|\omega_i|=\|\omega\|_\infty\}$

$$\|\partial_v^r\hat{h}_1\|_{1,\Lambda} = \|\chi_\lambda\partial_v^r\hat{h}_1\|_1 \le A(r,k,n,\mathcal{A})K_1^{n(k+1)}\left[n2^n\int_0^{K_1} K_1^{nN}y^{n-1}dy\right.$$

$$\left. + n2^n\int_{K_1}^{\lambda} y^N\prod_{j=1}^{n-1}\left[\int_0^{K_1}K_1^N dy_j + \int_{K_1}^{y} y_j^N dy_j\right]dy\right]$$

$$\le A(r,k,n,\mathcal{A})K_1^{n(k+1)}n2^n\ K_1^{n(N+1)}\left[\frac{1}{n} + \frac{1}{n}\left[\left(\frac{N+(\lambda/K_1)^{N+1}}{N+1}\right)^n - 1\right]\right] .$$

Note that, regardless of the sign of N,

$$\left[\frac{N+(\lambda/K_1)^{N+1}}{N+1}\right] \le \left[\frac{N+(\ell\pi\ )^{N+1}}{N+1}\right] .$$

∎

From Lemma 27 $B_i(r,k,n,\lambda)$ can be chosen to be independent of λ. For if $k+1 > 2r+3+1$ the $|\partial_v^r \hat{h}_i|$ is in $L^1(\mathbb{R}^n)$. Explicitly,

29. COROLLARY. For $k > 2r + 3$, (i.e., $N < -1$)

$$\|\partial_v^r \hat{h}_i\|_{1,\Lambda} \le \|\partial_v^r \hat{h}_i\|_1 \le A(r,k,n,\mathcal{A})K_1^{\ 2n(r+2)}2^n \left[\frac{(-N)}{(-N)-1}\right]^n .$$

In Lemma 22 it was seen that r could be as large as $k+1$. However, r must be less than half of this value for the Corollary to apply. For example, the Corollary does not apply for $k\le5$. For $k=9$, r can be no greater than 2. In this case $\eta_2(x_j)$ decreases as $\lambda^{-13/2}$ whereas $\eta_1(x_j)$ decreases as β^{-1} for $n = 2$. To have $\eta_1(x_j)$ converge more rapidly we must choose between large values for k, and hence for K_1, and a bound for $\eta_1(x_j)$ which depends on λ.

We can finally state our bound for $|\eta_1(x_j)|$. We conclude this section by collecting the results in

30. THEOREM. For E a compact subset of $\mathbb{R}^n$, for $\varphi = \varphi_{<k>s}$ as defined above, for $f \in L^1 \cap L^2(\mathbb{R}^n)$, for h_i, χ_λ, $\chi_{\mathcal{G}+\mathcal{M}}$, ψ_j, $\{x_j\}_{j\in J}$, and

μ_1 as defined above, and for $k \geq 3$,

$\varepsilon_3 =$

$$\left\|\chi_E\left(\sum_{j\in J}(\varphi*f)(x_j)\psi_j - \sum_{j\in J}\sum_{i=0}^{n}<\left(\left(\hat{h}_i\chi_\lambda\right)^\vee\right)_{[x_j]},\ \left(\chi_{\varphi+\mathcal{M}}\right)_{[x_j]}(\mu_i*f)>\psi_j\right)\right\|_p$$

$$\leq \left\|\chi_E\sum_{j\in J}\eta_1(x_j)\psi_j\right\|_p + \left\|\chi_E\sum_{j\in J}\eta_2(x_j)\psi_j\right\|_p$$

$$\leq \max_j\left\{|\eta_1(x_j)|\right\}\|\chi_E\|_p + \max_j\left\{|\eta_2(x_j)|\right\}\|\chi_E\|_p \;.$$

Let $C = \max\limits_{e=0,1,\ldots,r}\left\{\left[\left(\frac{\pi}{M}\right)^{4(e+1)}\left(\frac{m}{2}\right)^{2e+1}\right]^{1/(2e+3)}\right\}$, and let

$$A(r,k,n,\mathcal{A}) = ((r+1)!)^2\binom{k+r}{r}\left(\frac{5M^4}{8}\right)^{2n}\left(\frac{2}{m}\right)^n\left[\frac{s}{(k+1)C^{2n}} + \sqrt{n}M(n+1)r\left(\frac{5M^4}{8}\,\frac{2}{m}\right)^{2n}\right]^r$$

with $\mathcal{A}$ the set $\{a_0,a_1,\ldots,a_n)$ of which M and m are the max and min, respectively. Let $K_1 = \max\left\{2\frac{k+1}{s}\ ,\ C\right\}$, $\lambda = \ell\pi2\frac{k+1}{s}$, $\ell \in \mathbb{N}-\{0\}$. For $2r \geq n+1$ and for $k-2r-2\neq1$,

$$\max_j\{|\eta_1(x_j)|\} \leq (n+1)\|f\|_2\frac{1}{(2\pi)^n}\frac{\sqrt{n}}{\sqrt{2r-n}}\,2^{r+n}\,A(r,k,n,\mathcal{A})K_1^{2n(r+2)}\left(\frac{1}{\beta}\right)^{\frac{2r-n}{2}}$$

$$\times\left[1 + u\left(\left(\frac{N+(\ell\pi)^{N+1}}{N+1}\right)^n-1\right)\right],$$

and where $N = 2r+2-k$, $u = \begin{cases}0 \text{ if } \lambda\leq K_1\\ 1 \text{ if } \lambda>K_1\end{cases}.$

Secondly, for the case $2\frac{k+1}{s} \geq \left(\left(\frac{\pi}{M}\right)^4 \frac{m}{2}\right)^{1/3}$,

$$\max_j\{|\eta_2(x_j)|\}$$

$$\leq (n+1)\|f\|_2 \left(\frac{5M^4}{8}\right)^{2n} \left(\frac{2}{m}\right)^n \left(2\frac{k+1}{s}\right)^{\frac{7n}{2}} \left(\frac{n}{\pi^n}\right)^{\frac{1}{2}} \times$$

$$\left[\left(1+\frac{1+\ (\ell\pi)^{1+2(2-k)}}{2(k-2)\ -\ 1}\right)^{n-1} \frac{(\ell\pi)^{1+2(2-k)}}{2(k-2)-1}\right]^{\frac{1}{2}}.$$

3.7 Discrete approximate reconstruction

We have in this section the payoff for all of the preceding analysis: We can exhibit maps defined on discrete spaces which may be used in a digital implementation of the approximate reconstruction. For these maps we develop the final error term ε_4.

To begin, recall the interpolating function ψ used in the Construction section and in the Interpolation section: $\psi:\mathbb{R}^n \longrightarrow \mathbb{R}$ along with a discrete set of points $\mathcal{J}$ in $\mathbb{R}^n$ with index set J, $\{x_j\}_{j\in J} = \mathcal{J}$, such that, with $\psi_j = \psi_{[x_j]}$, $\chi_E = \chi_E \sum_{j\in J} \psi_j$, where E is a subset of $\mathbb{R}^n$ with compact closure. To condense the notation of the previous section, let

$$H_i = (\chi_\lambda \hat{h}_i)^{\vee}:\mathbb{R}^n \longrightarrow \mathbb{R}\ ,\quad i=0,1,\ldots,n\ ,$$

(recall $\hat{h}_i$ is symmetric) and let the set $\mathcal{S}+\mathcal{M}$ be denoted by $\mathcal{B}$. As in the construction section. let

$$G = \sum_{i=0}^{n} (H_i\chi_{\mathcal{B}})*\mu_i*f\ .$$

In the preceding sections we have developed the manner in which $\sum_{j\in J} G(x_j)\psi_j$ is an approximate reconstruction of f. The set $\mathcal{J}$ may be viewed as the "reconstruction set" in $\mathbb{R}^n$.

A second discrete subset of $\mathbb{R}^n$ is the "data set" $\mathcal{Q}$, the set on which the convolutions μ_i*f, $i=0,1,\ldots,n$, are evaluated. As in the Construction section let Q be the index set for $\mathcal{Q}$, $\mathcal{Q} = \{x_q\}_{q\in Q}$. We shall require

$$Q \supset \mathcal{J}$$

and

$$\text{for every } x_q \in Q,\ x_j \in \mathcal{J},\ q \in Q,\ j \in J : x_q = x_j \Leftrightarrow q = j .$$

With this notation the objective of this section is to exhibit a map

$$\tilde{H}_i : Q \cap \mathcal{B} \longrightarrow \mathbb{R}$$

such that the discrete convolution

$$\tilde{G}: \mathcal{J} \longrightarrow \mathbb{R}$$

$$\tilde{G}(x_j) = \sum_{i=0}^{n} \sum_{q\in Q} \left(\tilde{H}_i \chi_{\mathcal{B}}\right)(x_j - x_q) \ \left(\mu_i * f\right)(x_q)$$

approximates G in the sense that

$$\varepsilon_4 = \Big\| \chi_E \sum_{j\in J} \left(G(x_j) - \tilde{G}(x_j)\right) \psi_j \Big\|_p \xrightarrow[|Q| \to 0]{} 0 ,$$

where $|Q|$ is a suitable measure of the "mesh" of Q. Here the irregular notation $\tilde{H}_i \chi_{\mathcal{B}}$ is used in place of $\tilde{H}_i \left(\chi_{\mathcal{B}}|_Q\right)$. Also $\tilde{G}$ depends on Q, but this dependence is supressed in the notation. We have immediately

$$\varepsilon_4 \le \sum_{i=0}^{n} \max_{j\in J} \left\{ \left| \left[\left(H_i \chi_{\mathcal{B}}\right) * \mu_i * f \right](x_j) - \sum_{q\in Q} \left(\tilde{H}_i \chi_{\mathcal{B}}\right)(x_j - x_q) \ \left(\mu_i * f\right)(x_q) \right| \right\} \|\chi_E\|_p .$$

We require that the set Q have associated with it a set $S \subset \mathbb{R}^n$. With the notation $\chi_{S_q}(x) = \chi_S(x - x_q)$ for $q \in Q$, $x_q \in Q$, $x \in \mathbb{R}^n$, and for χ_S the characteristic function of S, the sets Q and S are to satisfy

i. $\chi_E * \chi_{\mathcal{B}} \sum_{q\in Q} \chi_{S_q} = \chi_E * \chi_{\mathcal{B}}$ almost everywhere, and

$\chi_{\mathcal{B}} = \sum_{x_q \in \mathcal{B}\cap Q} (\chi_S)_{[x_q]}$ almost everywhere

ii. for $x_j \in \mathcal{J}$, for $x_q, x_{q'} \in Q$,

$$\left(\chi_{S_q} * \chi_{S_{q'}}\right)(x_j) = \begin{cases} \|\chi_S\|_1 & \text{if } x_{q'} = x_j - x_q \\ 0 & \text{otherwise} \end{cases}$$

iii. $\chi_S(-x) = \chi_S(x)$, $x \in \mathbb{R}^n$.

With these conditions on S and Q the difference in the expression for ε_4 splits:

$$\left|\left[\left(H_1\chi_{\mathcal{B}}\right)*\mu_1*f\right](x_j) - \sum_{q\in Q}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_j-x_q)\ \left(\mu_1*f\right)(x_q)\right| \qquad (*)$$

$$= \left|\left[\left(\sum_{q'\in Q}\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\ \left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\ \chi_{S_{q'}}\right) *\mu_1*f\right](x_j)\right.$$

$$\left. + \left[\left(\sum_{q'\in Q}\frac{1}{\|\chi_S\|_1}\ \left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\ \chi_{S_{q'}}\right) * \left(\sum_{q\in Q}\left[\mu_1*f - \left(\mu_1*f\right)(x_q)\right]\chi_{S_q}\right)\right](x_q)\right| .$$

To define $\tilde{H}_1$ and to bound ε_4 we specify certain remaining choices. In particular, let the index set Q be a finite subset of $\mathbb{Z}^n$ and let $\hat{Q}$ be a second finite subset of $\mathbb{Z}^n$. Choose $\delta > 0$ and $\Delta > 0$ and let

$$Q = \left\{x_q = q\delta\ ,\ q\in Q\right\}, \qquad S = \overbrace{\left[-\tfrac{\delta}{2},\tfrac{\delta}{2}\right] \times \dots \times \left[-\tfrac{\delta}{2},\tfrac{\delta}{2}\right]}^{n \text{ times}} ,$$

$$\hat{Q} = \left\{v_t = t\Delta\ ,\ t\in\hat{Q}\right\}, \qquad \hat{S} = \overbrace{\left[-\tfrac{\Delta}{2},\tfrac{\Delta}{2}\right] \times \dots \times \left[-\tfrac{\Delta}{2},\tfrac{\Delta}{2}\right]}^{n \text{ times}} ,$$

where Δ and $\hat{Q}$ are chosen such that, with $\chi_{\hat{S}}(\omega - t\Delta) = \chi_{\hat{S}_t}(\omega)$, $\omega \in \mathbb{R}^n$,

$$\sum_{t\in\hat{Q}} \chi_{\hat{S}_t} = \chi_\lambda .$$

We now define

$$\frac{1}{\|\chi_S\|_1}\,\big(\tilde{H}_i\chi_{\mathcal{B}}\big)(x_q) = \Big[\sum_{t\in\hat{Q}} \big(\chi_\lambda \hat{h}_i\big)(t\Delta)\ \chi_{\hat{S}_t}\Big]^{\vee}(x_q)\ \chi_{\mathcal{B}}(x_q)$$

$$= \sum_{t\in\hat{Q}} \big(\chi_\lambda \hat{h}_i\big)(t\Delta)\ e^{i\Delta\delta(t\cdot q)}\ \big(\chi_{\hat{S}}\big)^{\vee}(q\delta)\ \chi_{\mathcal{B}}(q\delta)\ .$$

31. THEOREM. Let $\mu_i * f$, $i=0,1,\ldots,n$, be given on the set Q. where δ and Q are such that condition i holds for given sets E and $\mathcal{B}$. Let $\tilde{H}_i$, Δ, $\hat{Q}$, and H_i be as above. Let $\mathcal{J}$ be a subset of Q, with J the corresponding subset of the index set Q. such that as above $\chi_E = \chi_E \sum_{j\in J}\psi_j$. Then

$$\varepsilon_4 = \Big\|\chi_E \sum_{j\in J}\sum_{i=0}^{n}\Big[\Big[\big(H_i\chi_{\mathcal{B}}\big)*\mu_i*f\Big](x_j) - \sum_{q\in Q}\big(\tilde{H}_i\chi_{\mathcal{B}}\big)(x_j-x_q)\,\big(\mu_i*f\big)(x_q)\Big]\psi_j\Big\|_p$$

$$\le \|\chi_E\|_p\|f\|_1\ \Big[\frac{\lambda}{\pi}\Big]^n \times$$

$$\Bigg[\max_{\omega\in\Lambda}\Bigg\{A(1,k,n,\mathcal{A})\Big[2\frac{k+1}{s}\Big]^{n(k+1)}\frac{\prod_{j=1}^{n}\max\{C\ ,\ |\omega_j|\}^5}{\prod_{j=1}^{n}\max\Big\{2\frac{k+1}{s},\,|\omega_j|\Big\}^{k+1}}\Bigg\}\ \frac{\sqrt{n}\Delta}{2}\Big(1+\frac{n\lambda\delta}{2}\Big)$$

$$+\Big[\frac{5M^4}{8}\Big]^{2n}\Big[\frac{2}{m}\Big]^n\ K^{3n}\ \Big[\frac{n\lambda\delta}{2} + \|\chi_{\mathcal{B}}\|_1 m^{-n}\min\Big\{\frac{n\delta}{m},1\Big\}\Big]\Bigg]\ ,$$

where $\Lambda = \{\|\omega\|_\infty<\lambda\}$, $A(1,k,n,\mathcal{A})$ and C are as in Lemma 27, and where K is as in the proof of Propostion 16.

Proof. In the following let $1\le r,r'\le\infty$, with $1/r + 1/r' = 1$ and $1/\infty = 0$. A bound for the first term in the splitting (*) is

$$\left|\left[\left(\sum_{q'\in Q}\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\chi_{S_{q'}}\right) *\mu_1*f\right](x_j)\right|$$

$$\le \left\|\left(\sum_{q'\in Q}\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\chi_{S_{q'}}\right) *\mu_1*f\right\|_\infty$$

$$\le \left\|\sum_{q'\in Q}\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\chi_{S_{q'}}\right\|_r \|\mu_1\|_1 \|f\|_{r'} .$$

A bound for the first factor of this bounding product is

$$\left\|\sum_{q'\in Q}\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\chi_{S_{q'}}\right\|_r$$

$$\le \max_{q'\in Q}\left\{\left\|\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_{q'})\right]\chi_{S_{q'}}\right\|_\infty\right\}\|\chi_{\mathcal{B}}\|_r .$$

We now apply the definition of $\tilde{H}_1$ to bound

$$\left\|\left[H_1\chi_{\mathcal{B}} - \frac{1}{\|\chi_S\|_1}\left(\tilde{H}_1\chi_{\mathcal{B}}\right)(x_q)\right]\chi_{S_q}\right\|_\infty$$

$$\le \sup_{x\in S_q}\left\{\frac{1}{(2\pi)^n}\sum_{t\in\hat{Q}}\int_{\hat{S}_t}\left|\left(\chi_\lambda\hat{h}_1\right)(\omega)e^{i\omega\cdot x} - \left(\chi_\lambda\hat{h}_1\right)(t\Delta)e^{i\omega\cdot x_q}\right|d\omega\right\}$$

$$\le \sup_{x\in S_q}\left\{\frac{1}{(2\pi)^n}\sum_{t\in\hat{Q}}\left[\int_{\hat{S}_t}\left|\chi_\lambda\hat{h}_1 - \chi_\lambda\hat{h}_1(t\Delta)\right|d\omega\right.\right.$$

$$+ \int_{\hat{S}_t}\left|\left(\chi_\lambda\hat{h}_1\right)(\omega) - \left(\chi_\lambda\hat{h}_1\right)(t\Delta)\right|\left|e^{i\omega\cdot x} - e^{i\omega\cdot x_q}\right|d\omega$$

$$\left.\left.+ \int_{\hat{S}_t}\left|\left(\chi_\lambda\hat{h}_1\right)(\omega)\right|\left|e^{i\omega\cdot x} - e^{i\omega\cdot x_q}\right|d\omega\right]\right\} .$$

For $x \in S_q$ and $\|\omega\|_\infty < \lambda$,

$$\left|e^{i\omega\cdot x} - e^{i\omega\cdot x_q}\right| \le |\omega\cdot(x-x_q)| \le |\omega|\,|(x-x_q)| \le \frac{n\lambda\delta}{2} .$$

Combining these bounds

$$\max_{i,q}\left\{\left\| \left[H_i\chi_{\mathcal{B}} - \frac{1}{\|\chi_{S}\|_1}\left(\tilde{H}_i\chi_{\mathcal{B}}\right)(x_q)\right]\chi_{S_q}\right\|_\infty\right\}$$

$$\le \frac{1}{(2\pi)^n}\max_i\left\{\max_{t\in\hat{Q}}\sup_{\omega\in\hat{S}_t}\left\{\left|\left(\chi_\lambda\hat{h}_i\right)(\omega) - \left(\chi_\lambda\hat{h}_i\right)(t\Delta)\right|\right\}\|\chi_\lambda\|_1\left(1+\frac{n\lambda\delta}{2}\right)\right.$$

$$\left. + \frac{n\lambda\delta}{2}\|\chi_\lambda\hat{h}_i\|_1\right\}.$$

We have $\|\chi_\lambda\|_1 = (2\lambda)^n$, and from the proof of Proposition 16

$$|\chi_\lambda\hat{h}_i\|_1 \le \|\chi_\lambda\|_1\ \|\hat{h}_i\|_\infty \le (2\lambda)^n\left[\frac{5M^4}{8}\right]^{2n}\left[\frac{2}{m}\right]^n K^{n(k+1)}K^{(2-k)n},$$

where, as always, $k\ge3$.

To bound the sup above we observe that with $v = \omega - t\Delta$ there exists $\omega' \in \hat{S}_t$ such that

$$\left|\left(\chi_\lambda\hat{h}_i\right)(\omega) - \left(\chi_\lambda\hat{h}_i\right)(t\Delta)\right| \le \left|\partial_v\hat{h}_i(\omega')\right| \le \max_{\omega'\in\Lambda}\left\{\left|\partial_{v/|v|}\hat{h}_i(\omega')\right|\right\}|\omega - t\Delta|$$

$$\le \max_{\omega'\in\Lambda}\left\{\left|\partial_{v/|v|}\hat{h}_i(\omega')\right|\right\}\sqrt{n}\,\frac{\Delta}{2}.$$

By Lemma 27

$$\sup_{\omega\in\hat{S}_t}\left\{\left|\left(\chi_\lambda\hat{h}_i\right)(\omega) - \left(\chi_\lambda\hat{h}_i\right)(t\Delta)\right|\right\}$$

$$\le \max_{\omega\in\Lambda}\left\{A(1,k,n,\mathcal{A})\left[2\frac{k+1}{s}\right]^{n(k+1)}\frac{\prod_{j=1}^{n}\max\{C,\ |\omega_j|\}^5}{\prod_{j=1}^{n}\max\left\{2\frac{k+1}{s},|\omega_j|\right\}^{k+1}}\right\}\frac{\sqrt{n}\Delta}{2}.$$

This completes the bound for the first term of the splitting (*).

For the second term of the splitting we first use

$$\left|\left[\left(\sum_{q'\in Q}\frac{1}{\|\chi_S\|_1}\left(\tilde{H}_i\chi_{\mathcal{B}}\right)(x_{q'})\,\chi_{S_{q'}}\right) * \left(\sum_{q\in Q}\left[\mu_i * f - \left(\mu_i * f\right)(x_q)\right]\chi_{S_q}\right)\right](x_q)\right|$$

$$\le \Big\| \sum_{q'\in Q}\frac{1}{\|\chi_S\|_1}\left(\tilde{H}_i\chi_{\mathcal{B}}\right)(x_{q'})\,\chi_{S_{q'}}\Big\|_r\ \Big\|\sum_{q\in Q}\left[\mu_i * f - \left(\mu_i * f\right)(x_q)\right]\chi_{S_q}\Big\|_{r'}\ .$$

To bound $\|\sum_{q\in Q}\left[\mu_i * f - \left(\mu_i * f\right)(x_q)\right]\chi_{S_q}\|_{r'}$ note that, with $1\le v, v'\le\infty$, $1/v + 1/v' = 1$,

$$\Big|\sum_{q\in Q}\left[\mu_i * f(x) - \left(\mu_i * f\right)(x_q)\right]\chi_{S_q}(x)\Big|$$

$$\le \sum_{q\in Q}\|\left(\mu_i\right)_{[x]} - \left(\mu_i\right)_{[x_q]}\|_v\ \|f\|_{v'}\chi_{S_q}(x)$$

$$\le \left(\frac{1}{a_i}\right)^{n-n/v}\min\left\{\frac{n\delta}{a_i}\ ,\ \left(\frac{n\delta}{a_i}\right)^{1/v},\ 2^{1/v}\right\}\|f\|_{v'}\sum_{q\in Q}\chi_{S_q}(x)\ ,$$

where $\delta/2 = \max_{x\in S_q}\left\{\|x - x_q\|_\infty\right\}$. The last inequality follows from Lemma 10 and from $\varphi_{<0>a_i} = \mu_i$. As usual, all occurrences of a_i in the last expression may be replace by m.

To bound $\|\sum_{q'\in Q}\frac{1}{\|\chi_S\|_1}\left(\tilde{H}_i\chi_{\mathcal{B}}\right)(x_{q'})\,\chi_{S_{q'}}\|_r$ use

$$\max_{i,q}\left\{\Big|\frac{1}{\|\chi_S\|_1}\left(\tilde{H}_i\chi_{\mathcal{B}}\right)(x_{q'})\Big| \le \max_i\left\{\frac{1}{(2\pi)^n}\ \Big\|\sum_{t\in\hat{Q}}\left(\chi_\lambda\hat{h}_i\right)(t\Delta)\chi_{\hat{S}_t}\Big\|_1\right\}\right.$$

$$\le \max_i\left\{\|\chi_\lambda\hat{h}_i\|_\infty\right\}\frac{\|\chi_\lambda\|_1}{(2\pi)^n} \le \left(\frac{\lambda}{\pi}\right)^n\|\hat{h}_i\|_\infty$$

$$\le \left(\frac{\lambda}{\pi}\right)^n\left(\frac{5M^4}{8}\right)^{2n}\left(\frac{2}{m}\right)^n K^{3n}.$$

To complete the proof, combine the above bounds and use r′=1 in the

bounds for the first term of the splitting $(*)$, and use $r = v' = 1$ in the bounds for the second term of the splitting.

3.8 Discussion

A primary motivation for exhibiting an explicit error bound was to determine if a "practical" support for the deconvolutors $(\chi_\lambda \hat{h}_1)^\vee \chi_{\mathcal{B}}$ could be established. A "practical" support would be one for which the side length β of $\mathcal{B}$ differed from the side length m of the smallest convolutor by a factor of several tens. Such a support would be useful for applications.

The bounds established here do not satisfy our "practical" criterion. Let us examine the error ε for a specific case. Consider

$$n = 2, \qquad \mathcal{A} = \{1,\ \sqrt{2},\ \sqrt{3}\}, \qquad k \geq 3, \qquad \text{and} \qquad s \leq 1.$$

Then

$$m = 1, \qquad M = \sqrt{3}, \qquad 2\frac{k+1}{s} \geq 8, \ \text{and}$$

(see the Theorem in Approximate reconstruction for definitions)

$$C = \left(\left(\frac{\pi}{M}\right)^4 \frac{m}{2} \right)^{1/3} \cong 1.76, \qquad K_1 = 2\frac{k+1}{s}\ .$$

Since the side length $m = 1$ of the smallest convolutor is our unit in $\mathbb{R}$, it is easy to select a function f and a set E such that

$$\|f\|_p \leq 1, \quad 1 \leq p \leq \infty, \qquad \text{and} \quad \|\chi_E\|_1 = 2^n$$

(e.g., a simple function with support in E). For such a case the error ε should be no more than 1.

Consider ε_3. In the bound for ε_3 given in Section 6 the quantities $|\eta_1(x_j)|$ and $|\eta_2(x_j)|$ have the common factor

$$L = (n+1)\sqrt{n}\left(\frac{5M^4}{8}\right)^{2n}\left(\frac{2}{m}\right)^n \cong 1.70 \times 10^4 .$$

This term also appears in $A(r,k,n,\mathscr{A})$ so that

$$|\eta_1(x_j)| \le \|f\|_2 \frac{2^{r+n}}{(2\pi)^n\sqrt{2r-n}}\,((r+1)!)^2 \binom{k+r}{r}\left[1+rM2^n L\right]^r L\, K_1^{4n+2nr}\left(\frac{1}{\beta}\right)^{\frac{2r-n}{2}},$$

$$|\eta_2(x_j)| \le \|f\|_2 \left(\frac{1}{\pi}\right)^{n/2} 3^{(n-1)/2}\, L\, K_1^{3n+n/2}\left(\frac{1}{\ell\pi}\right)^{k-3+1/2}.$$

If $k = 3$, then for $\varepsilon \le \|f\|_2$ it is necessary that $|\eta_2(x_j)| \le \|f\|_2$ which requires that

$$\ell\pi \ge K_1^{7n} \ge 8^{14} = 2^{42} .$$

For $|\eta_1(x_j)|$ to not exceed $|\eta_2(x_j)|$ it is necessary that

$$\beta^{r-n/2} \ge (\ell\pi)^{1/2}\, K_1^{n(2r+1/2)},$$

whence, for $r = 2$,

$$\beta \ge (\ell\pi)^{1/2}\, 8^9 \ge 2^{21+27} = 2^{48}.$$

Clearly, such estimates are not "practical." Similar relations hold for Δ and δ that appear in the bound for ε_4.

4 MULTIPLE OPERATOR DECONVOLUTION WITH ADDITIVE NOISE; THE ENVELOPE OPERATOR

SUMMARY

The methods for multiple operator deconvolution of Berenstein, Taylor, and Yger are examined for the case of the addition of a noise signal after each of the multiple convolutions and preceding the deconvolutions. It is shown that for strongly coprime multiple operators there is an obvious choice for optimal deconvolvers. The case of m strongly coprime, parallel convolvers with m independent noise sources is compared to that of m identical, parallel convolvers with m independent, identically distributed noise sources. A performance criterion is defined. The performance for selected collections of strongly coprime convolvers is shown to be at least as good as that for the corresponding collection of an equal number of identical, parallel convolvers. That is, there is no penalty for the additional frequency response available with deconvolution, at least for the noncompactly supported optimal deconvolvers. Qualitative methods are developed to characterize the properties of strongly coprime configurations. These methods enable the description of circumstances in which it is advantageous to use strongly coprime multiple detectors of large support.

4.1 Introduction

Throughout the last several years mathematical results have been presented which form the foundations for the use of multiple (parallel) linear operators, each given by convolution with a distinct kernel (or impulse response), in place of the use of a single such linear operator or, equivalently, in place of the use of multiple (parallel) operators each with the identical kernel (Kelleher and Taylor 1971; Berenstein and Taylor 1979, 1980a, 1980b; Berenstein, Taylor, and Yger 1983a, 1983b; Berenstein and Yger 1983; Berenstein 1983). See Figure 16. In the multiple operator method each distinct kernel (also referred to as a convolver or convolutor) is associated with a second kernel, referred to as a deconvolver. These kernels are viewed as distributions, that is, as linear functionals on the space of infinitely differentiable functions on $\mathbb{R}^n$. The mathematical results cited above describe the conditions under which compactly supported distributions $\mu_1, \mu_2, \ldots, \mu_m$ have associated to them compactly supported distributions $\nu_1, \nu_2, \ldots, \nu_m$ such that

$$\sum_{i=1}^{m} \mu_i * \nu_i = \delta , \tag{1}$$

where δ is the Dirac distribution on $\mathbb{R}^n$ and where $*$ denotes convolution.

This is of interest for applications in which the convolver μ_i must correspond to a physical, analog device wherein the impulse

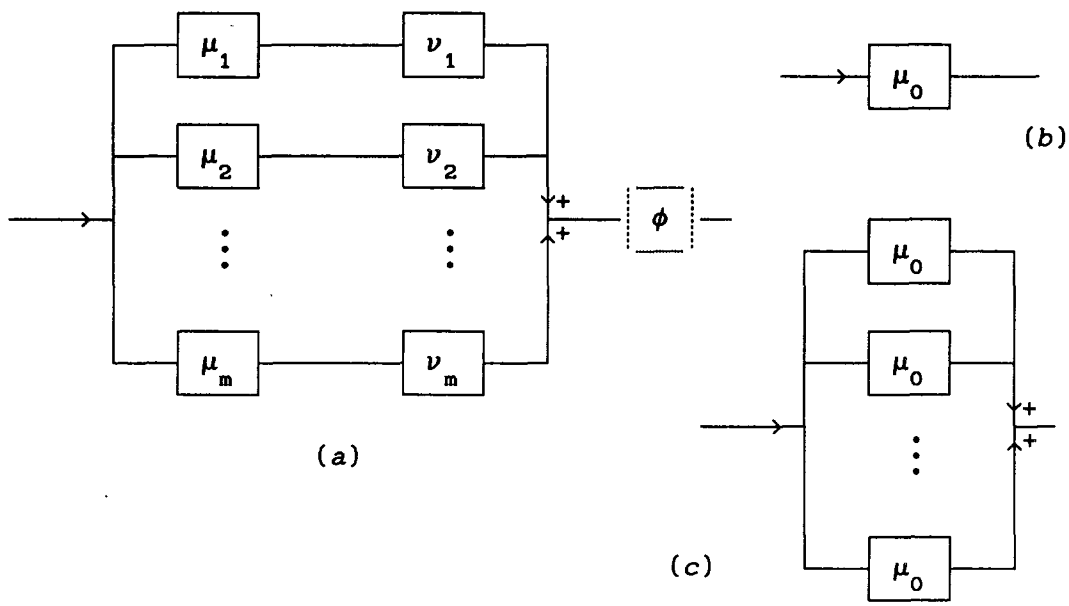


Fig. 16. (*a*) Multiple parallel linear operators with distinct distributions μ_i. Single operator (*b*) and multiple parallel operators with identical distributions μ_0 (*c*).

response is dictated by a solid state or biological process. It is entirely possible to select such analog convolvers which satisfy approximately the multiple operator criteria. Then each associated deconvolver can be digitally implemented. The fact that the deconvolvers are linear and of compact support means that their implementation is straightforward; that they are continuous implies

stability. Most importantly, the evident high bandwidth of the overall operator is accomplished without any essential change in the response functions of the analog devices. The term overall operator refers to the operator given by the kernel distribution $\sum_{i=1}^{m} \mu_i * \nu_i = \delta$. Of course, because of practical constraints such as analog and digital approximations and computation time, the design objective for the overall operator would not be the identity operator with impulse response δ but rather a high bandwidth approximation of the identity operator given by an impulse response ϕ. In terms of the distribution equation (1) and since convolutions commute

$$\sum_{i=1}^{m} (\mu_i * \phi) * \nu_i = \sum_{i=1}^{m} \mu_i * (\nu_i * \phi) = \phi . \tag{2}$$

In a sense ϕ can be considered to be made up of "parts," each of which arises from one of the practical constraints just listed, along with a special part that is deliberately added to control the noise power spectrum of the output of the overall operator.

The publications on this subject have appeared primarily in the mathematical literature. The following issues regarding (1) have been addressed: sufficient conditions for the existence of solutions (Hörmander 1967; Kelleher and Taylor 1971; Berenstein and Taylor 1979, 1980a, 1980b); examples of sets of distributions that satisfy the sufficient conditions (Berenstein, Taylor, and Yger 1983a, 1983b); construction of explicit solutions, that is, explicit formulas for the deconvolvers (Berenstein and Yger 1983; Berenstein 1983); and construction and evaluation of approximate solutions (Berenstein,

Krishnaprasad, and Taylor 1984; Chapter 3 of this document).

Only recently have specific applications of (1) been mentioned. The work of Berenstein, Krishnaprasad, and Taylor (1984) addressed the case of one dimensional integration over an interval as a linear operator on a variety of function spaces. This work was the first time that (1) and contemporary mathematical methods for understanding the equation were applied to physical problems. There the linear operators in (1) were considered to act on function spaces other than the space of infinitely differentiable functions $C^{\infty}(\mathbb{R}^{n})$. In applications these other function spaces may be $L^{p}(\mathbb{R}^{n})$ (functions with modulus to the power p having bounded Lebesgue integral) or, more generally, Sobolev spaces. The consideration of (1) acting on such functions spaces requires the consideration of $C^{\infty}(\mathbb{R}^{n})$ as a dense subset and the behavior of the operators on the closure. Consequently it is natural that approximate identities and mollifiers such as ϕ in (2) are used. This work also discussed the question of additive noise and the question of the continuity of the overall operator with respect to the distributions μ_1, $\mu_2, \ldots,$ μ_m. The noise question is in regard to noise added following the action of the operators defined by the μ_i, while the continuity question is in regard to the dependence of the overall performance on either the actual analog approximations of the μ_i or the digital approximations of the ν_i .

The approximation methods of Chapter 3 of this document were motivated by this work of Berenstein et al. These methods exploit the

approximation in (2). In conjunction with the analysis of the methods (Chapter 3), a computer simulation for $\mathbb{R}^2$ was performed. This simulation dramatically illustrated (2) for imaging devices in which the analog convolvers were solid state photodetectors. With these results there was an increased interest in imaging applications. This led to the consideration of not just detectors but of linear systems consisting of sequences of operators with each operator of the multiple operator type. These activities led to the need to answer basic systems analysis questions.

This chapter describes the result of our application of standard methods of linear systems and random signals to the multiple operator type of system of equations (1) and (2). This analysis was necessary if one was to seriously consider multiple operator designs. While the extended bandwidth was well understood, analyzed, and even illustrated in simulations, the consequence of the introduction of noise and of design errors was not fully understood. It was clear that since the operator was linear and continuous that there would be no instablity due to noise (at least for smooth ($C^{\infty}(\mathbb{R}^n)$) approximations), which is already an improvement over the case of single operator reconstruction methods (Berenstein, Taylor, and Yger 1983a; Berenstein, Krishnaprasad, and Taylor 1984). However, the performance needed to be explicity described so that standard tools such as resolution, equivalent bandwidth, and signal to noise ratio would be available for systems engineering design studies.

This investigation was motivated in large part by the potential application of these multiple operator methods to electo-optics, especially to imaging devices. We have in mind imaging devices that are for the detection, transformation, and display of electromagnetic radiation for a human observer as well as such devices for artificially intelligent "observers." Consequently, the problems and the desired solutions have the flavor of this application. While the analysis and the results are in a sense general, much is framed and guided by the motivating problems.

With this in mind, let us review two features of performance descriptions suitable for engineering studies. For imaging electro-optics systems it is best to cast off any hope and preference, common in mathematics, for an obvious choice of norm or metric as a performance measure. First, performance criteria are never uniquely determined by the device: they depend instead on the infinite number of possible end-uses. Loosely speaking, if there are two end-uses that are "linearly independent," then one would need at least either two real valued performance metrics or a performance criterion that takes values in a 2-dimensional space. For example, for field use of an infrared imaging device for observations in a natural terrain, there is a requirement for good sensitivity at low spatial frequencies for purposes of orientation and search strategy relative to the terrain, while there is a requirement for sufficient response at sufficiently high spatial frequencies for purposes of accomplishing the objective of

the observation (Ratches, Lawson, et al. 1975). These two sub-uses of field use are an example of two "independent" uses. A different end-use, say industrial robot vision, would surely have distinct sub-uses that were independent of those in the field use example.

The simplest thing to hope for is a performance criteria that can be "projected" onto any of the criteria "spanned" by a set of end-uses. Consequently, it is typical in electo-optics to use functions to characterize devices and systems and to rarely be satisfied with a choice of norm of the function, or even with a choice of a projection of the function to a finite dimensional space. In other words, one is willing to forego a linear ordering of devices.

The transfer function (the Fourier transform of the impulse response) and the noise power spectral density are familiar examples of such device characterizing functions. (The second may depend on a background signal level as well as the device.) The minimum resolvable temperature difference is another such function. (A human observer is assumed for this one.) The simiplest example of a projection is the "evaluate at" map; for example, evaluating the modulus of the transfer function at a specific frequency projects the space of all transfer functions onto the real numbers. Evaluating a weighted sum of such projections is a further example. While the set of all transfer functions in not naturally ordered, it inherits an order from a fixed choice of such a projection, as well as from, say, the L^2 norm. The characterization of a device by such dimension reducing projections is

almost always inadequate; one prefers to see the system in terms of its characterizing functions.

On the other hand, system characterizations typically have an implied equivalence relation. A familiar example of such an equivalence relation is the one in which transfer functions that differ by a constant, nonzero multiple are identified (hence the use of the familiar signal to noise ratio). That is, by means of a suitable equivalence relation one seeks to factor from the characterizing functions all irrelevant differences. (In the example, any difference in gain is to be neglected.) Frequently the equivalence classes are identified by a standard choice of normalization. (In optics, transfer functions are normalized to unity at zero frequency.) The identification and use of equivalence classes reduces the size of the space of the characterizing functions.

The objective of this chapter is to provide explicit performance characterizations for multiple operator deconvolution in the presence of additive noise. In addition to the two features above (functions, equivalence classes) that are to be incorporated, a third is that characterizations are always relative: the whole point of any characterization is comparisons. Our objective, then, is to provide an explicit performance characterization for multiple operator deconvolution relative to the performance of any of the constituent single operators. Once this is accomplished the existing comparisons between conventional single operators can be used to compare multiple

operators with arbitrary single operators. And our goal is to do this with wisely chosen equivalence classes so that succint engineering conclusions can be formed directly from the characterizing functions. This goal is accomplished in this chapter by the use of what we call the *envelope operator* (and the equivalence class it generates) associated with a multiple operator. With this construction the comparison task is reduced to a comparison of transfer functions.

4.2 General results

A fundamental result in this subject is the following. Given a set of distributions $\mu_1, \mu_2,\dots,\mu_m$ on $\mathbb{R}^n$, each with compact support, then the necessary and sufficient condition for the existence of a second set of distributions $\nu_1, \nu_2,\dots,\nu_m$ on $\mathbb{R}^n$, again each with compact support, such that

$$\sum_{i=1}^{m} \mu_i * \nu_i = \delta , \tag{1}$$

is that the Fourier-Laplace transforms of the μ_i, denoted $\hat{\mu}_i$, satisfy

$$\sum_{i=1}^{m} |\hat{\mu}_i(z)| \geq C_1 e^{-C_2|\mathrm{Im}\, z|}(1+|z|)^{-N} , \quad z \in \mathbb{C}^n \tag{3}$$

for some positive constants C_1, C_2, and N (Hörmander 1967; Kelleher and Taylor 1971). (For $\zeta = (\zeta_1, \zeta_2,\dots,\zeta_n) \in \mathbb{C}^n$ define $|\zeta| = \left(\sum_i |\zeta_i|^2 \right)^{1/2}$.) The condition (3) is often referred to as the *strongly coprime* condition.

Here we will need only elementary harmonic analysis and we shall consider the Fourier transform on $\mathbb{R}^n$, that is, the restriction of the Fourier-Laplace transform to $\mathbb{R}^n \subset \mathbb{C}^n$ in the sense that for $z = (z_1, z_2,\dots,z_n) \in \mathbb{C}^n$, $\omega = (\omega_1, \omega_2,\dots, \omega_n) = (\mathrm{Re}\, z_1, \mathrm{Re}\, z_2,\dots, \mathrm{Re}\, z_n) \in \mathbb{R}^n$. Then (3) has the form

$$\sum_{i=1}^{m} |\hat{\mu}_i(\omega)| \geq C_1(1+|\omega|)^{-N} , \quad \omega \in \mathbb{R}^n . \tag{4}$$

For any distribution ν of compact support, $\hat{\nu} \in C^\infty(\mathbb{R}^n)$. As usual, we may choose $\phi \in C^\infty(\mathbb{R}^n)$ such that $\hat{\phi}$ has compact support and is

sufficiently differentiable so that $\nu*\phi \in L^1(\mathbb{R}^n)$. But $\nu*\phi$ can not have compact support. However, for each $i=1,2,\ldots m$ define $h_i = \nu_i*\phi \in L^1(\mathbb{R}^n)$. Then

$$\sum_i \mu_i*h_i = \phi. \tag{5}$$

The $h_i \in L^1(\mathbb{R}^n)$ that satisfy (5) are not uniquely determined. From (4) and from $\hat{\mu}_i \in C^\infty(\mathbb{R}^n)$ and with ϕ as above, the choice

$$D_i(\omega) = \frac{\bar{\hat{\mu}}_i(\omega)}{\sum_{j=1}^{m} |\hat{\mu}_j(\omega)|^2}, \quad \hat{h}_i(\omega) = D_i(\omega)\hat{\phi}(\omega), \quad i=1,2,\ldots,m, \tag{6}$$

defines functions $h_i \in L^1(\mathbb{R}^n)$ which satisfy (5). ($\bar{z}$ denotes the complex conjugate of z.)

While (6) is exhibited essentially by inspection, the result can be obtained in a more systematic fashion as well as in a more general form. We first recall some standard tools, apply these tools to a simple case, and then proceed to the more general form. The diagram in Figure 17 represents an operator L acting on a function f. Let (temporarily) f be bounded and in $C^\infty(\mathbb{R}^n)$. Let $\mu_1, \mu_2,\ldots,\mu_m$ be an arbitrary set of m distributions with compact support. For each linear operator defined by μ_i let η_i be a sample function of a zero mean, wide-sense stationary random process that is added to the output of μ_i, let $\eta_i \in L^\infty(\mathbb{R}^n)$, and let N_i^2 ($N_i \geq 0$) be the noise power spectral density of the process (see, for example, Davenport and Root 1958, Ch.4, Ch.6). For each distinct i and j let η_i be independent of η_j and let each η_j be independent of f. Let ν_i be defined by $(\nu_i*\phi)^\wedge = D_i\hat{\phi}$,

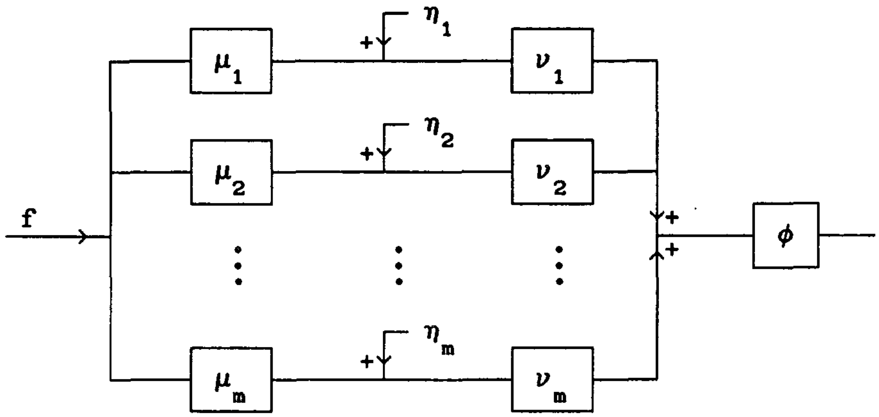


Fig. 17. Multiple operator configuration consisting of m parallel convolvers μ_i, m noise signals η_i, and m deconvolvers ν_i.

where $\hat{\phi}$, $D_i \in C^r(\mathbb{R}^n)$, $\hat{\phi}$ has compact support, and r is sufficiently large so that $\left[D_i\hat{\phi}\right]^\wedge \in L^1(\mathbb{R}^n)$. Let $g \in L^\infty(\mathbb{R}^n)$ be defined by

$$g = Lf = \sum_{i=1}^{m} (\mu_i * f + \eta_i)*(\nu_i * \phi). \tag{7}$$

In the usual manner, with E denoting expectation,

$$E\{g\} = \sum_{i=1}^{m} \mu_i * f * (\nu_i * \phi). \tag{8}$$

Let T'_y denote translation by y, $T'_y(x) = x + y$, let $^\vee$ denote inverse

Fourier transform, and let $\| \ \|_p$ denote the L^p norm. Directly from the definition of wide-sense stationary and noise power spectral density it follows that

$$E\left\{\left(g - E\{g\}\right)\left[\left(g - E\{g\}\right)\circ T'_y\right]\right\} = \left[\sum_{i=1}^{m} N_i^2 |D_i|^2 |\hat{\phi}|^2\right]^{\vee}(y) \ , \tag{9a}$$

and, for $y = 0$, that

$$E\left\{\left(g - E\{g\}\right)^2\right\} = \frac{1}{(2\pi)^n}\left\| \sum_{i=1}^{m} N_i^2 |D_i|^2 |\hat{\phi}|^2 \right\|_1 . \tag{9b}$$

The simplest configuration for L is all distributions equal, all deconvolvers trivial, and all random processes identically distributed:

$$\mu_i = \mu_0 \ , \quad \nu_i = \delta \ , \quad N_i^2 = N_0^2 \ , \quad \text{for } i = 1,2,\ldots,m \ . \tag{10}$$

Then

$$E\{g\} = m\mu_0 * \phi * f \ , \quad E\left\{\left(g - E\{g\}\right)^2\right\} = \frac{m}{(2\pi)^n}\|N_0|\hat{\phi}|\|_2^2 \ . \tag{11}$$

The utility of (8) and (9) or of (11) is that if L is followed by a linear operator $\mathfrak{U}$ with kernel $\mathfrak{u}$ (which could model a specific "end-$\mathfrak{U}$se") then classical discrimination methods would compare the function $\left(\mathfrak{U}(E\{g\})\right)^2$ with the constant function $E\left\{\left(\mathfrak{U}\left(g-E\{g\}\right)\right)^2\right\}$. In the case of the simplest configuration, (10) and (11), there are the following formulas and bounds.

$$\left(\mathfrak{U}(E\{g\})\right)^2 = E\{\mathfrak{U}g\}^2 = \left(\mathfrak{u}*(m\mu_0*\phi*f)\right)^2 = m^2\left[\left(\hat{\mathfrak{u}}\hat{\mu}_0\hat{\phi}\hat{f}\right)^{\vee}\right]^2 \tag{12}$$

$$\le \left[\frac{m}{(2\pi)^n}\ \|\hat{\mathfrak{u}}\hat{\mu}_0\hat{\phi}\hat{f}\|_1\right]^2$$

$$\le \left[\frac{m}{(2\pi)^n}\right]^2 \|\hat{\mathfrak{u}}\hat{\mu}_0\hat{\phi}\|_2^2 \ \|\hat{f}\|_2^2 \qquad \text{when } f \in L^2(\mathbb{R}^n) \ ;$$

and

$$E\left\{\left(\mathfrak{U}\left(g - E\{g\}\right)\right)^2\right\} = \frac{m}{(2\pi)^n}\|\hat{\mathfrak{u}}N_0\hat{\phi}\|_2^2 . \tag{13}$$

The function $E\{\mathfrak{U}g\}$ is referred to as the signal, its square $E\{\mathfrak{U}g\}^2$ is referred to as the signal power or energy, and $E\left\{\left(\mathfrak{U}\left(g-E\{g\}\right)\right)^2\right\}$ is referred to as the noise power. Typically the ratio of $E\{\mathfrak{U}g\}^2$ to $E\left\{\left(\mathfrak{U}\left(g-E\{g\}\right)\right)^2\right\}$ is considered, or, alternatively, the positive square root of the ratio. Here we shall consistently use the latter. If this ratio is evaluated at some distinguished point, the value defines a "signal to noise ratio." We denote by $\mathfrak{P}$ the projection of a function by the evaluation of the absolute value of the function at the distinguished point. Given L and for a given choice of ϕ, f, $\mathfrak{U}$, and $\mathfrak{P}$ define the signal to noise ratio

$$\mathcal{SNR}(\mathfrak{U}L) = \frac{\mathfrak{P}\mathfrak{U}(E\{g\})}{\left[E\left\{\left(\mathfrak{U}\left(g-E\{g\}\right)\right)^2\right\}\right]^{1/2}} . \tag{14}$$

For a fixed choice of ϕ, f, $\mathfrak{U}$, and $\mathfrak{P}$, two operators L and L′ can be compared and ordered by (14).

On the other hand, for a choice of ϕ, f, $\mathfrak{U}$, and $\mathfrak{P}$, (14) is determined for the case of the trivial operator in (10) by the pair of functions

$$m\,\hat{\mu}_0 \text{ and } \sqrt{m}\,N_0 . \tag{15}$$

In general, let operators L and L′ (for example, as in Figure 17) have transfer functions and noise power spectral densities $\hat{\mu}$, N^2 and $\hat{\mu}'$, N'^2, respectively. For a choice of $\mathfrak{U}$ we shall say that $\mathfrak{U}L|\mathfrak{U}L'$

(i.e., "$\mathfrak{U}L$ divides $\mathfrak{U}L'$") if there exists a function $\hat{q} \in L^{\infty}(\mathbb{R}^n)$ such that $\hat{\mathfrak{u}}\,\hat{q}\,\hat{\mu} = \hat{\mathfrak{u}}\,\hat{\mu}'$. If $\mathfrak{U}L | \mathfrak{U}L'$ and $|\hat{q}|^2|\hat{\mathfrak{u}}|^2N^2 \le |\hat{\mathfrak{u}}|^2 N'^2$, we say that $\mathfrak{U}L \ge \mathfrak{U}L'$.

This definition is motivated by the following. As usual, let ϕ be such that a linear operator Ω with kernel q can be associated with $\hat{q}$ by considering $\hat{q}\hat{\phi}$. Let $\mathfrak{B}$ be any continuous, translation invariant, linear operator. For fixed $\mathfrak{U}$ if $\mathfrak{U}L \ge \mathfrak{U}L'$, then $\dfrac{\mathcal{SNR}(\mathfrak{U}\mathfrak{B}\Omega L)}{\mathcal{SNR}(\mathfrak{U}\mathfrak{B}L')} \ge 1$.

Consequently, $\sup\limits_{\mathfrak{B}} \mathcal{SNR}(\mathfrak{U}\mathfrak{B}L) \ge \sup\limits_{\mathfrak{B}} \mathcal{SNR}(\mathfrak{U}\mathfrak{B}L')$.

Next consider the operator L diagrammed in Figure 17 for the case in which $\mu_1,\ \mu_2,\ldots,\mu_m$ are distinct and strongly coprime (i.e., satisfy (3)). An obvious consequence is $\sum\limits_{i=1}^{m} |\hat{\mu}_i(\omega)|^2 > 0$ and, equivalently,

$$0 \ne (\hat{\mu}_1(\omega),\ \hat{\mu}_2(\omega),\ldots,\hat{\mu}_m(\omega)) \in \mathbb{C}^n\ ,\quad \omega \in \mathbb{R}^n\ . \tag{16}$$

Consequently we can visualize (16) as is shown in Figure 18*a*. A similar illustration can be used to visualize $\hat{f}(\omega)(\hat{\mu}_1(\omega),\ \hat{\mu}_2(\omega),\ldots,\hat{\mu}_m(\omega)) = (\hat{f}(\omega)\hat{\mu}_1(\omega),\ \hat{f}(\omega)\hat{\mu}_2(\omega),\ldots,\hat{f}(\omega)\hat{\mu}_m(\omega))$, except the "curve" passes through the origin if and only if $\hat{f}(\omega) = 0$. The power spectral densities are real and nonnegative (thus we write N_i^2 and choose $N_i \ge 0$). Assume

$$N_i(\omega) > 0\ ,\quad \omega \in \mathbb{R}^n\ ,\quad i = 1,2,\ldots,m\ . \tag{17}$$

We can visualize (17) as is shown in Figure 18*b*. The case of strongly coprime multiple operators has the useful feature that the consideration of (16) and (17) pointwise in conjuction with (8) and (9)

uniquely determines an alternative choice for the D_i of (6). This choice will be optimal in the sense it has the smallest $E\left\{\left(g-E\{g\}\right)^2\right\}$ among all sets of deconvolvers.

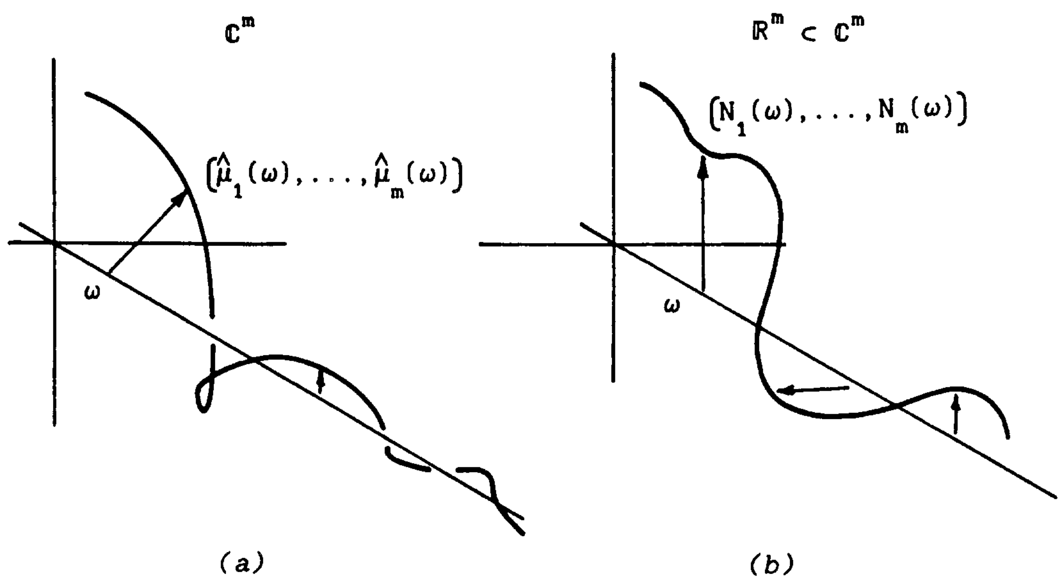


Figure 18

PROPOSITION. For $N_i \in L^\infty(\mathbb{R}^n)$, $N_i(\omega) > 0$ for $\omega \in \mathbb{R}^n$, $i = 1,2,\ldots,m$, then $D : \mathbb{R}^n \longrightarrow \mathbb{C}^m$ is uniquely determined (almost everywhere) by the conditions, for fixed $\omega \in \mathbb{R}^n$,

$$D(\omega) = (D_1(\omega),\ D_2(\omega),\ldots,D_m(\omega)) = z$$

minimizes $\sum_{i=1}^{m} |z_i|^2 N_i^2(\omega)$ on the set $\left\{z \in \mathbb{C}^m : \sum_{i=1}^{m} z_i\hat{\mu}_i(\omega) = 1\right\}$. (18)

In fact

$$D_i(\omega) = \frac{\dfrac{\overline{\hat{\mu}_i(\omega)}}{N_i^2(\omega)}}{\displaystyle\sum_{i=1}^{m} \frac{|\hat{\mu}_j(\omega)|^2}{N_i^2(\omega)}} \tag{19}$$

Proof: Any z that satisfies (18) is clearly contained in the linear subspace of $\mathbb{C}^n$ determined by the span of

$$\left\{ (\hat{\mu}_1(\omega),0,\ldots,0),\ (0,\hat{\mu}_2(\omega),0,\ldots,0),\ldots,\ (0,0,\ldots 0,\hat{\mu}_m(\omega)) \right\}. \tag{20}$$

That is, $z_i = 0$ if $\hat{\mu}_i(\omega) = 0$. Equivalently, there exists $\lambda = (\lambda_1, \lambda_2,\ldots,\lambda_m) \in \mathbb{C}^m$ such that

$$\left(z_1N_1(\omega),\ z_2N_2(\omega),\ldots,z_mN_m(\omega)\right) = \left(\lambda_1\overline{\hat{\mu}_1}(\omega),\ \lambda_2\overline{\hat{\mu}_2}(\omega),\ldots,\lambda_m\overline{\hat{\mu}_m}(\omega)\right). \tag{21}$$

Let $\sum_i'$ denote $\displaystyle\sum_{\substack{i=1\\ \hat{\mu}_i(\omega)\neq 0}}^{m}$. Then (18) implies

$$\text{minimize} \quad \sum_i' |\lambda_i|^2|\hat{\mu}_i(\omega)|^2 \quad \text{on} \quad \left\{ \sum_i'\lambda_i \frac{|\hat{\mu}_i(\omega)|^2}{N_i(\omega)} = 1 \right\}. \tag{22}$$

From this it follows that the λ_i are all real, so that (22) in the form

$$\text{minimize} \quad \sum_i' \left(\lambda_i|\hat{\mu}_i(\omega)|\right)^2 \quad \text{on} \quad \left\{ \sum_i'\lambda_i|\hat{\mu}_i(\omega)| \ \frac{|\hat{\mu}_i(\omega)|}{N_i(\omega)} = 1 \right\} \tag{23}$$

is an elementary case for $\mathbb{R}^n$ and has the unique solution

$$\lambda_i |\hat{\mu}_i(\omega)| = \frac{\dfrac{|\hat{\mu}_i(\omega)|}{N_i(\omega)}}{\displaystyle\sum_{j=1}^{m} \frac{|\hat{\mu}_j(\omega)|^2}{N_j^2(\omega)}} \quad (\text{for } \hat{\mu}_i(\omega) \neq 0) \ . \tag{24}$$

Consequently, from (21), the unique z corresponding to the minimum is $D(\omega)$ as in (19). ■

In addition to $N_i > 0$, $i = 1,2,\dots,m$, we shall assume $N_0 > 0$. Further, we shall assume that the N_i are sufficiently differentiable and that $\frac{1}{N_i} = \mathcal{O}(|\omega|^p)$ for some integer p, $i = 0,1,2,\dots,m$. With this we can find $\hat{\phi} = \mathcal{O}(|\omega|^{-p'})$ so that $(D_i\hat{\phi})^{\vee} \in L^2(\mathbb{R}^n)$ and for $\hat{\phi}$ sufficiently smooth and with compact support then $(D_i\hat{\phi})^{\vee} \in L^1(\mathbb{R}^n)$.

COROLLARY. For the choice of D_i from the Proposition,

$$\sum_{i=1}^{m} \hat{\mu}_i D_i = 1 \quad , \quad \left(\sum_{i=1}^{m} |D_i|^2 N_i^2 \right)^{1/2} = \left[\frac{1}{\displaystyle\sum_{j=1}^{m} \frac{|\hat{\mu}_j|^2}{N_j^2}} \right]^{1/2} . \tag{25}$$

Let L_0 identify the trivial configuration of L in (10) and let L_s identify the strongly coprime configuration. Unless explicitly indicated to the contrary, L_s indicates that the deconvolvers D_i of (19) are used. The first of the functions in (25) is the transfer function for L_s and the second is the square root of the noise power

spectral density. The corresponding functions for L_0 are (15). The mollifier $\hat{\phi}$ is suppressed but understood. From (25) obviously $L_s | L$ for any operator L. (Note that L always denotes a pair, a transfer function (linear operator) and an additive noise.) From (15) and (25) the dividend $\hat{q}$ for $L = L_0$ is $m\hat{\mu}_0$. Let N_s^2 denote the noise power spectral density of L_s. In the sense discussed earlier let Ω_0 denote the linear operator associated with $m\hat{\mu}_0$. That $L_s | L_0$ with dividend $m\hat{\mu}_0$ means $L_0 = \Omega_0 L_s$. Then $\Omega_0 L_s$ has functions corresponding to (25) (transfer function, square root of noise power spectral density) given by

$$m\hat{\mu}_0 \sum_{i=1}^{m} \hat{\mu}_i D_i = m\hat{\mu}_0 ,$$

$$|m\hat{\mu}_0| N_s = m|\hat{\mu}_0| \left[\sum_{i=1}^{m} |D_i|^2 N_i^2 \right]^{1/2} = \left[\frac{m \dfrac{|\hat{\mu}_0|^2}{N_0^2}}{\displaystyle\sum_{j=1}^{m} \frac{|\hat{\mu}_j|^2}{N_j^2}} \right]^{1/2} \sqrt{m} N_0 \tag{26}$$

By definition $\mathfrak{U}L_s \geq \mathfrak{U}L_0$ if $|m\hat{\mu}_0| N_s(\omega) \leq \sqrt{m} N_0(\omega)$ on the support of $\hat{u}$, and $\mathfrak{U}L_0 \geq \mathfrak{U}L_s$ if $\mathfrak{U}L_0 | \mathfrak{U}L_s$ and $|m\hat{\mu}_0| N_s(\omega) \geq \sqrt{m} N_0(\omega)$ on the support of $\hat{u}$. Thus, whether $\mathfrak{U}L_0 \geq \mathfrak{U}L_s$ or $\mathfrak{U}L_s \geq \mathfrak{U}L_0$ holds depends, in part, on whether one of the following inequalities holds on the support of $\hat{u}$: from (15) and (26)

$$|m\hat{\mu}_0| N_s(\omega) \ \begin{smallmatrix} \leq \\ \geq \end{smallmatrix} \ \sqrt{m} N_0(\omega) \quad \Longleftrightarrow$$

$$\left[\sum_{i=1}^{m} \frac{N_0^2(\omega)}{N_i^2(\omega)} \, |\hat{\mu}_i(\omega)|^2 \right]^{1/2} \ \begin{smallmatrix} \geq \\ \leq \end{smallmatrix} \ \sqrt{m} |\hat{\mu}_0(\omega)| \, . \tag{27}$$

In (27) the notation means that the upper inequality symbol on the left is to be paired with the upper inequality symbol on the right and lower left with lower right.

The comparison in (27) can in special cases be viewed from a slightly different perspective. First, view the left side of the second inequality in (27) as the Fourier transform of a kernel. Define

$$\hat{e}(\omega) = \left[\sum_{i=1}^{m} \frac{N_0^2(\omega)}{N_i^2(\omega)} \, |\hat{\mu}_i(\omega)|^2 \right]^{1/2} . \tag{28}$$

We refer to $\hat{e}$ as the *envelope* transfer function corresponding to the *envelope operator* $\mathfrak{E}$ for a given strongly coprime L_s in comparison with a given L_0. If $\sqrt{m}\mathfrak{E}$ acts on L_s, then the pair of functions associated with $\sqrt{m}\mathfrak{E}L_s$ is

$$\sqrt{m}\,\hat{e} \quad , \quad \sqrt{m}\,N_0 \, . \tag{29}$$

Recall that the pair for L_0 is given by (15) (rewritten for convenience)

$$m\,\hat{\mu}_0 \quad , \quad \sqrt{m}\,N_0 \, . \tag{15}$$

That is, the composition of $\sqrt{m}\mathfrak{E}$ with L_s has a noise power spectral density equal to that of L_0. If, for example, $\hat{\mu}_0$ is real and positive, then it makes sense to compare (29) with (15). It is easy to check that the condition $\hat{\mathfrak{u}}\widehat{\sqrt{m}e} \geq \hat{\mathfrak{u}}m\hat{\mu}_0$ (on the support of $\hat{\mathfrak{u}}$) coincides with our definition $\mathfrak{u}L_s \geq \mathfrak{u}L_0$, and $\hat{\mathfrak{u}}m\hat{\mu}_0 \geq \hat{\mathfrak{u}}\widehat{\sqrt{m}e}$ coincides with what we mean by $\mathfrak{u}L_0 \geq \mathfrak{u}L_s$. These two inequalities are precisely the content of the comparison of the right side of (27). One could say that $\sqrt{m}\mathfrak{E}$ is the normalization of L_s to the noise power spectral density of L_0.

For either point of view we consider $W_> = \left\{\omega \in \mathbb{R}^n\colon \hat{e}(\omega) \geq \sqrt{m}|\hat{\mu}_0(\omega)|\right\}$ and $W_< = \left\{\omega \in \mathbb{R}^n\colon \hat{e}(\omega) \leq \sqrt{m}|\hat{\mu}_0(\omega)|\right\}$. For all $\mathfrak{u}$ such that $\hat{\mathfrak{u}}$ has support in $W_>$ it follows from (27) and the definitions that $\mathfrak{u}L_s \geq \mathfrak{u}L_0$. Consequently,

$$\frac{\mathcal{SNR}(\mathfrak{u}\Omega_0 L_s)}{\mathcal{SNR}(\mathfrak{u}L_0)} = \frac{\| \hat{\mathfrak{u}}\, \hat{\phi}\, N_0 \|_2}{\left\| \hat{\mathfrak{u}}\, \dfrac{\sqrt{m}\, \hat{\mu}_0}{\hat{e}}\, \hat{\phi}\, N_0 \right\|_2} \geq 1 , \tag{30}$$

where Ω_0 is used to denote the linear operator corresponding to the transfer function $m\hat{\mu}_0$ of L_0 .

Assume $\hat{\mu}_0(0) \neq 0$ and define

$$\Omega_0 = \left\{ \omega \in \mathbb{R}^n : \forall t \in [0,1) \ |\hat{\mu}_0(t\omega)| > 0 \right\} .$$

Note that for $\mathbb{R}^1$ the usual definition of limiting resolution is $\sup \Omega_0$. If $\operatorname{supp}(\hat{\mathfrak{u}})$ is compact and $\operatorname{supp}(\hat{\mathfrak{u}}) \subset \Omega_0$, then $\mathfrak{u}\Omega_0^{-1}$ makes sense, consequently $\mathfrak{u}L_0 | \mathfrak{u}L_s$. Hence, if $\operatorname{supp}(\hat{\mathfrak{u}})$ is compact and $\operatorname{supp}(\hat{\mathfrak{u}}) \subset W_< \cap \Omega_0$, then $\mathfrak{u}L_0 \geq \mathfrak{u}L_s$. Consequently,

$$\frac{\mathcal{SNR}(\mathfrak{u}L_s)}{\mathcal{SNR}(\mathfrak{u}\Omega_0^{-1}L_0)} = \frac{\left\| \hat{\mathfrak{u}}\, \dfrac{1}{\sqrt{m}\, \hat{\mu}_0}\, \hat{\phi}\, N_0 \right\|_2}{\left\| \hat{\mathfrak{u}}\, \dfrac{1}{\hat{e}}\, \hat{\phi}\, N_0 \right\|_2} \leq 1 . \tag{31}$$

In general the inequality cannot be extended to all of $W_< \cap \Omega_0$ because of the behavior of $1/\hat{\mu}_0$ on the boundary.

There is no information regarding $\dfrac{\mathcal{SNR}(\mathfrak{u}L_s)}{\mathcal{SNR}(\mathfrak{u}L_0)}$ implied by either $\mathfrak{u}L_s \geq \mathfrak{u}L_0$ or $\mathfrak{u}L_0 \geq \mathfrak{u}L_s$. Additional information is needed. For example, it may be sufficient to know the effect of the so-called "boost" $\mathfrak{u}L_0 \longmapsto \mathfrak{u}\Omega_0^{-1}L_0$. In particular, if $\mathrm{supp}(\hat{\mathfrak{u}})$ is compact and $\mathrm{supp}(\hat{\mathfrak{u}})\subset\Omega_0$ then

$$\mathrm{supp}(\hat{\mathfrak{u}}) \subset W_> \quad\text{and}\quad \frac{\mathcal{SNR}(\mathfrak{u}\Omega_0^{-1}L_0)}{\mathcal{SNR}(\mathfrak{u}L_0)} \geq 1 \quad\Longrightarrow\quad \frac{\mathcal{SNR}(\mathfrak{u}L_s)}{\mathcal{SNR}(\mathfrak{u}L_0)} \geq 1\ , \tag{32a}$$

and

$$\mathrm{supp}(\hat{\mathfrak{u}}) \subset W_< \quad\text{and}\quad \frac{\mathcal{SNR}(\mathfrak{u}\Omega_0^{-1}L_0)}{\mathcal{SNR}(\mathfrak{u}L_0)} \leq 1 \quad\Longrightarrow\quad \frac{\mathcal{SNR}(\mathfrak{u}L_s)}{\mathcal{SNR}(\mathfrak{u}L_0)} \leq 1\ . \tag{32b}$$

For $\mathrm{supp}(\hat{\mathfrak{u}}) \subset \mathbb{R}^n-\Omega_0$, or even for $\mathrm{supp}(\hat{\mathfrak{u}})\cap(\mathbb{R}^n-\Omega_0) \neq \emptyset$, it is often the case in applications that $\mathfrak{u}L_0$ is not defined. Since $\mathfrak{u}L_s$ is defined for all $\mathfrak{u}$ it makes sense in such cases to consider $\mathfrak{u}L_s \geq \mathfrak{u}L_0$.

4.3 Examples: Characteristic functions of sets in R^n

Collections of sets in $\mathbb{R}^n$ such that the characteristic functions of the sets in the collection are strongly coprime have been reported (Berenstein, Taylor, and Yger 1983a, 1983b; Chapter 3 herein). For example, such a collection of cubes consists of $m = n+1$ cubes in $\mathbb{R}^n$ with sides parallel and with side lengths $\sqrt{a_1}$, $\sqrt{a_2}, \ldots,$ $\sqrt{a_m}$, where for all $i \neq j$ a_i and a_j are relatively prime integers and for all i $\sqrt{a_i}$ is not an integer. A second example is the collection of $m = 2$ disks in $\mathbb{R}^2$ where the ratio of the radii is an integer between 2 and 200.

A common situation for electro-optic detectors on $\mathbb{R}^n$ [e.g., n=1 (slits), n=2 (focal plane arrays), n=3 ($\mathbb{R}^2 \times$\{time\})] is for the noise power spectral density to have the form $\|\chi_S\|_1 N_\varnothing^2$, where $\|\chi_S\|_1$ is the L^1 norm of the characteristic function χ_S of the set S (equivalently, the Lebesgue measure of the set). For such a case, let sets S_1, $S_2, \ldots,$ S_m, be chosen so that, for $\mu_i = \chi_{S_i}$, the μ_1, $\mu_2, \ldots, \mu_m$ are strongly coprime. Then, from the Proposition,

$$D_i(\omega) = \frac{\dfrac{\overline{\hat{\mu}_i(\omega)}}{\|\mu_i\|_1}}{\displaystyle\sum_{j=1}^{m} \frac{|\hat{\mu}_j(\omega)|^2}{\|\mu_j\|_1}} \tag{33}$$

and (25) becomes

$$\sum_{i=1}^{m} \hat{\mu}_i D_i = 1 \quad , \quad \left(\sum_{i=1}^{m} |D_i|^2 N_i^2 \right)^{1/2} = \left[\frac{1}{\sum\limits_{j=1}^{m} \frac{|\hat{\mu}_j|^2}{\|\mu_j\|_1}} \right]^{1/2} N_\emptyset \quad . \tag{34}$$

Let S_0 be any set, let $\mu_0 = \chi_{S_0}$ be its characteristic function, and consider this to be the convolver in L_0 defined by (10) (i.e., m parallel, identical convolvers). Let the noise power spectral density have the same form as above, $N_0^2 = \|\mu_0\|_1 N_\emptyset^2$. From (27) and (28) one obtains an envelope transfer function $\hat{e}_d$ and the associated comparison for these two: a convenient renormalization by the constant $\|\mu_0\|_1^{1/2}$ is made in

$$\hat{e}_d(\omega) = \frac{\hat{e}(\omega)}{\|\mu_0\|_1^{1/2}} = \left[\sum_{i=1}^{m} \frac{|\hat{\mu}_i(\omega)|^2}{\|\mu_i\|_1} \right]^{1/2} \gtrless \sqrt{m}\, \frac{|\hat{\mu}_0(\omega)|}{\|\mu_0\|_1^{1/2}} \quad . \tag{35}$$

For an explicit example let $S_i \subset \mathbb{R}^2$ be the region in a focal plane of an imaging device which corresponds to a single light sensitive detector. The exposure time interval is assumed fixed and the image is assumed constant. Then $\mu_i = \chi_{S_i}$ is the idealized response function of the detector. (The actual shape of the response function, if not deconvolved, is incorporated into the mollifier ϕ.) Then $\hat{\mu}_i$ is what is referred to as the "detector MTF" and the form of the noise power spectral density corresponds to typical detector properties such as "D^*" for infrared detectors. The density has the above form as well for the so-called background limited case. It also has this form for

$\mathbb{R}^3$ when the time interval is included as the third dimension. Further, a background limited slit detector corresponds to the above forms for $\mathbb{R}^1$ with the slit width as the coordinate. (In the background limited case there is assumed to be a relatively small signal of interest superimposed on a relatively large constant signal so that the noise in the signal of interest is due to the "shot" noise of the constant signal.) (For detector characteristics discussed above see, for example, Kingston 1978, Ch.2.)

In Figures 19 and 20 the transfer functions for such cases are shown. In Figure 19, a comparison is shown for the example for $\mathbb{R}^1$. The characteristic functions μ_1 and μ_2 for the two intervals $[-1\ ,\ 1]$ and $[-\sqrt{2}\ ,\ \sqrt{2}]$, respectively, are strongly coprime. The envelope transfer function $\hat{e}_d$ is shown and is compared with the transfer function for the two identical, parallel convolvers as in (35) where $\mu_0 = \mu_1$. The choice $\mu_0 = \mu_1$ is used rather than $\mu_0 = \mu_2$ in this comparison because μ_1 is "better" than μ_2 in the sense that the first zero of $\hat{\mu}_1$ (i.e., its bandwidth) is greater than the first zero of $\hat{\mu}_2$. Recall from the scaling property for Fourier transforms on $\mathbb{R}^1$ that $\mu_1(x) = \mu_2(\sqrt{2}x)$ for all $x \in \mathbb{R}^1$ if and only if $\sqrt{2}\hat{\mu}_1(\sqrt{2}\omega) = \hat{\mu}_2(\omega)$ for all $\omega \in \mathbb{R}^1$. Figure 19 illustrates the consequence of the strongly coprime condition: the envelope response is approximately an envelope for the modulus of the other two responses and, correspondingly, is without zeroes. Also, it can be observed that the envelope response decreases approximately as $1/|\omega|$.

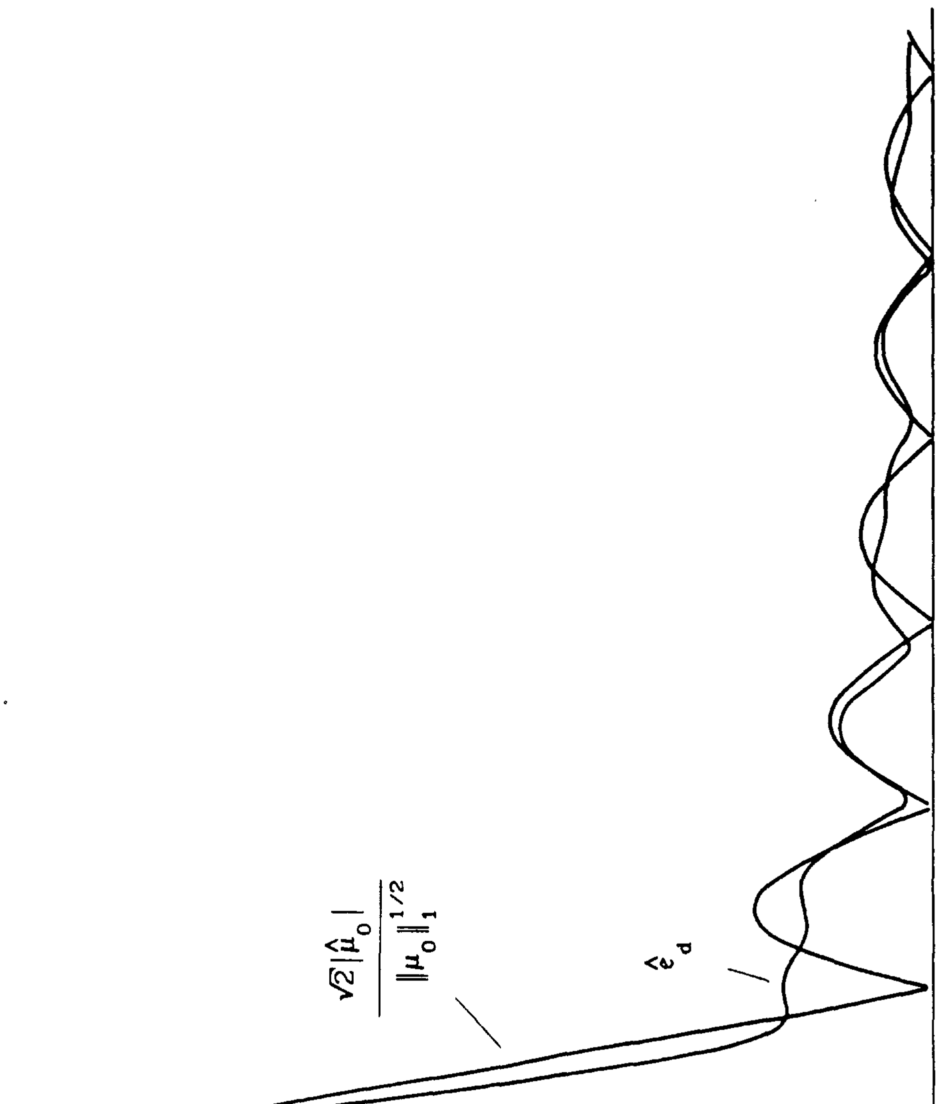


Fig. 19. $\mathbb{R}^1$

In Figure 20 the envelope transfer function is shown for an example in $\mathbb{R}^2$, the case of three squares Q_1, Q_2, Q_3 of side length 1, $\sqrt{2}$, $\sqrt{3}$, respectively. The characteristic functions of these three squares are strongly coprime. The comparison (35) is illustrated in Figure 20 by graphing the modulus of the corresponding transfer functions for two subsets of $\mathbb{R}^2$: the ω_1-axis $\left\{\omega=(\omega_1,\omega_2) \in \mathbb{R}^2 : \omega_2=0\right\}$ (see Figure 20*a*) and the diagonal $\left\{\omega=(\omega_1,\omega_2) \in \mathbb{R}^2 : \omega_1=\omega_2\right\}$ (see Figure 20*b*). All graphs use the Euclidean distance as abscissa, $|\omega| = \left(\omega_1^2 + \omega_2^2\right)^{1/2}$. The comparison illustrated in Figure 20 is for $\hat{\mu}_0 = \hat{\chi}_{Q_1}$. (As before, $\hat{\chi}_{Q_1}$ has the greatest bandwidth and the scaling property for $\mathbb{R}^n$ has the form $\mu_1(x) = \mu_2(kx)$ for $k > 0$ and for all $x \in \mathbb{R}^n$ if and only if $k^n\hat{\mu}_1(k\omega) = \hat{\mu}_2(\omega)$ for all $\omega \in \mathbb{R}^n$.) The comparison is essentially the same as that for the two intervals in $\mathbb{R}^1$. The difference between the ω_1-axis and the diagonal illustrates that approximately the envelope response decreases as $|\omega|^{-1}$ along the ω_1-axis and as $|\omega|^{-2}$ along the diagonal.

From (35) (and as illustrated by the figures) the following statements can be made. These are stated as "observations" because the results can not be given in terms of explicit inequalities. Some notation is helpful. Define

$$\Omega_i = \left\{ \omega \in \mathbb{R}^n : \forall t \in [0,1] \ |\mu_i(t\omega)| > 0 \right\} \quad \text{and} \quad \Omega = \bigcap_{i=1}^{m} \Omega_i \ . \tag{36}$$

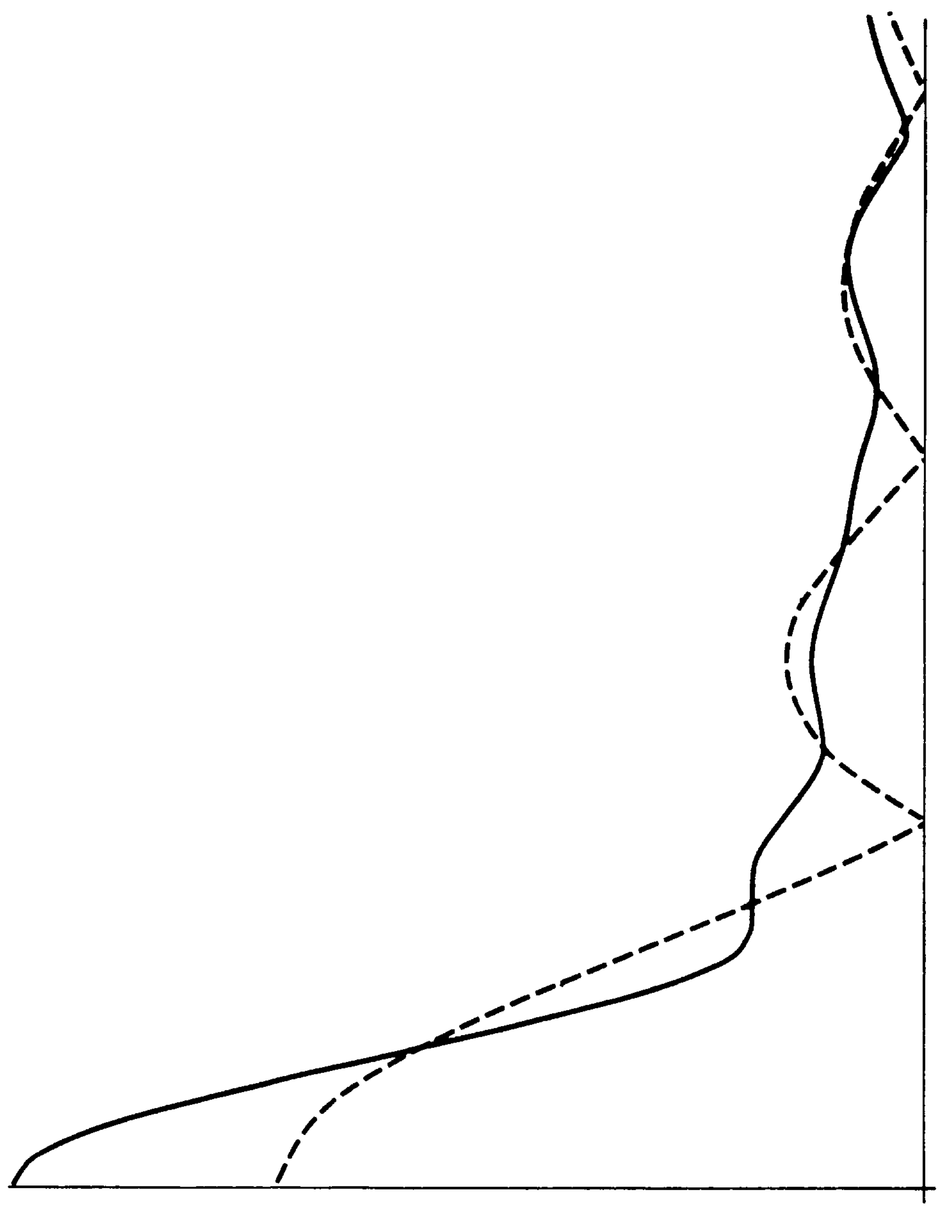


Fig. 20a. $\mathbb{R}^2$; abscissa = ω_1-axis solid = $\hat{\varepsilon}_d$; broken: $\sqrt{3}\,\dfrac{|\hat{\mu}_0|}{\|\mu_0\|_1^{1/2}}$, $\hat{\mu}_0 = \hat{\chi}_{O_1}$

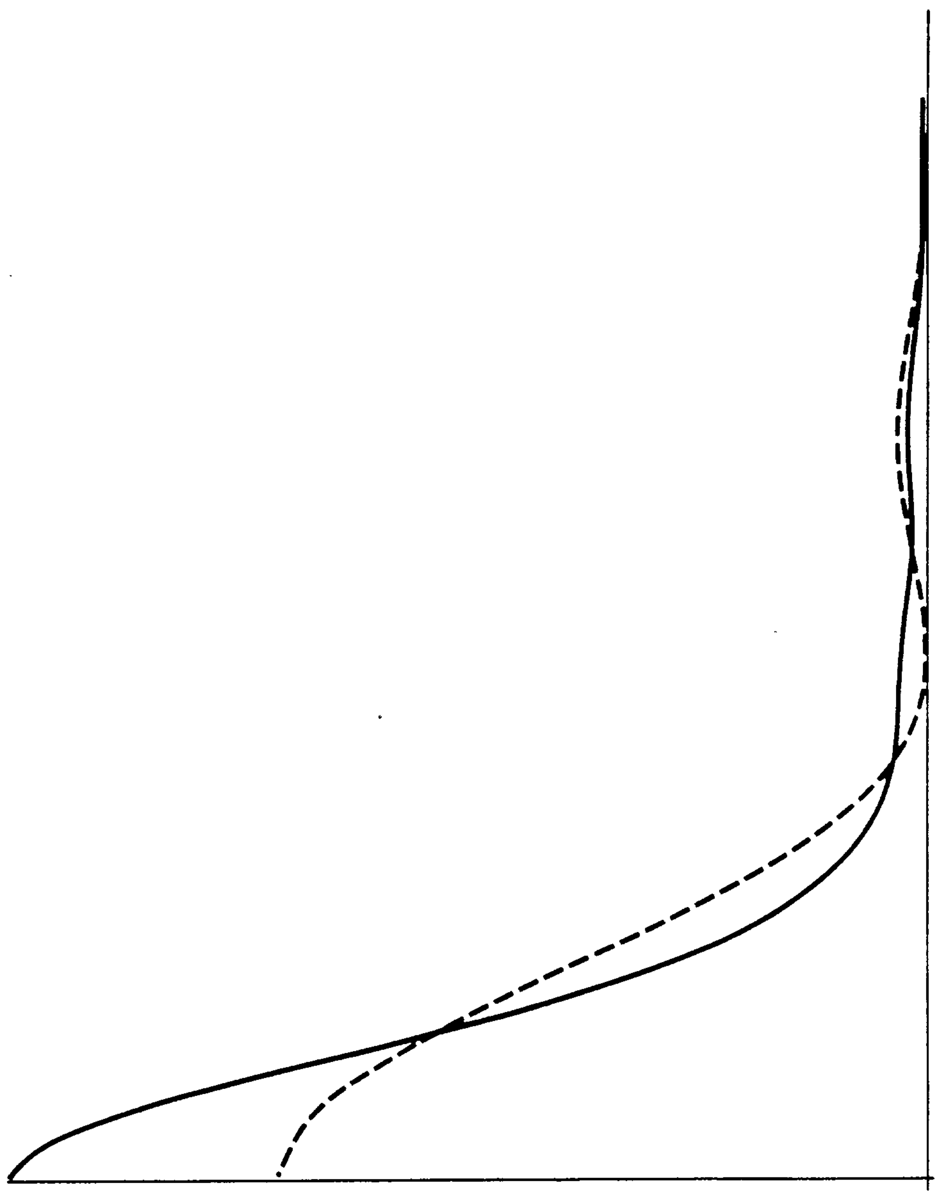


Fig. 20*b*. $\mathbb{R}^2$; abscissa = diagonal solid = $\hat{e}_d$; broken: $\sqrt{3}\,\dfrac{|\hat{\mu}_0|}{\|\mu_0\|_1^{1/2}}$, $\hat{\mu}_0 = \hat{\chi}_{0_1}$

OBSERVATIONS: Let $\mu_1, \mu_2, \ldots, \mu_m \in L^1(\mathbb{R}^n)$ be strongly coprime characteristic functions of sets in $\mathbb{R}^n$ as considered above. With each μ_i let there be associated as in Figure 17 an additive wide- sense stationary noise with noise power spectral density of the form $\|\mu_i\|_1 N_\emptyset^2$. Let L_s be the configuration in Figure 17 with deconvolvers determined by the Proposition. Let L_0 be the trivial configuration as in (10) with $\mu_0 = \mu_1$, $N_0 = N_1$.

Observation 1: For $\mathfrak{u}$ with $\mathrm{supp}(\hat{\mathfrak{u}}) \subset \Omega$, $\mathfrak{u}L_s \cong \mathfrak{u}L_0$.

Observation 2: For $\mathfrak{u}$ with $\mathrm{supp}(\hat{\mathfrak{u}}) \subset \mathbb{R}^n - \bigcup_{i=2}^{m} \Omega_i$, $\mathfrak{u}L_s \geq \mathfrak{u}L_0$.

Observation 3: For $\mathfrak{u}$ with $\mathrm{supp}(\hat{\mathfrak{u}}) \subset \bigcup_{i=2}^{m} \Omega_i - \Omega$, $\mathfrak{u}L_s \leq \mathfrak{u}L_0$.

Observation 4: For $\mathfrak{u}$ with $\mathrm{supp}(\hat{\mathfrak{u}})$ compact, $\mathrm{supp}(\hat{\mathfrak{u}}) \subset \Omega_1$, let Ω_0^{-1} be the boost on $\mathfrak{u}$, $\mathfrak{u}L_0 \longmapsto \mathfrak{u}\Omega_0^{-1}L_0$ (see (31) and (32)). If *Observation 3* can be neglected then

$$\frac{\mathcal{SNR}(\mathfrak{u}\Omega_0^{-1}L_0)}{\mathcal{SNR}(\mathfrak{u}L_0)} \geq 1 \quad\Longrightarrow\quad \frac{\mathcal{SNR}(\mathfrak{u}L_s)}{\mathcal{SNR}(\mathfrak{u}L_0)} \geq 1 .$$

As discussed at (31) it is not possible to extend this to all of Ω_1, for $\hat{\mu}_1 = 0$ on the boundary of Ω_1. However, it still is desirable to have a means to compare L_s with the more well known, more thoroughly studied trivial configurations. In the next section this is accomplished by pushing the troublesome set $\{\hat{\mu}_1 = 0\}$ out toward infinity.

AN EXPLICIT EXAMPLE

It is instructive to consider $\dfrac{\mathcal{SNR}(\mathfrak{U}L_s)}{\mathcal{SNR}(\mathfrak{U}L_0)}$ for some explicit choices for $N_\varnothing$, f, $\mathfrak{U}$, and for $\mathfrak{P}$. Let $N_\varnothing$ be the constant function. For fixed $\omega_0 \in \mathbb{R}^n$, let $f(x) = 1 + \exp(i\omega_0 \cdot x)$, $x \in \mathbb{R}^n$. Let $\mathfrak{P}\mathfrak{U}(E\{g\}) = \left|\left(\mathfrak{U}(E\{g\})\right)(0)\right|$, and let $\hat{u} = \hat{u}_\ell$ be the characteristic function of the punctured (excludes $\omega = 0$) closed disk centered at the origin and of radius ℓ, $|\omega_0| < \ell$. (Alternatively, the punctured cube of side lenght 2ℓ, or the set $\left\{\omega\in\mathbb{R}^n : |\omega_i| \le \ell_i, i=1,2,\ldots,n\right\}$. This latter alternative, $\hat{u}_\ell$ the characteristic function of the set $\left\{\omega \in \mathbb{R}^2 : |\omega_1| \le 2|\omega_0|, |\omega_2| \le \frac{2}{5}|\omega_0|\right\}$, with $f(x) = 1 + \exp\left(ix_1(\omega_0)_1\right)$, is a very coarse approximation of a standard vision model (Ratches, Lawson, et al. 1975).) Then from (14), (34), and (35), with C a constant,

$$\mathcal{SNR}(\mathfrak{U}L_s) = C\,\frac{|\hat{\phi}(\omega_0)\hat{u}(\omega_0)|}{\left\| \dfrac{\hat{\phi}\,\hat{u}\,N_\varnothing}{\hat{e}_d} \right\|_2} \qquad \mathcal{SNR}(\mathfrak{U}L_0) = C\,\frac{|\hat{\phi}(\omega_0)\,\sqrt{m}\,\hat{\mu}_1(\omega_0)\,\hat{u}(\omega_0)|}{\|\mu_1\|_1^{1/2}\,\|\hat{\phi}\,\hat{u}\,N_\varnothing\|_2}$$

hence

$$\frac{\mathcal{SNR}(\mathfrak{U}L_s)}{\mathcal{SNR}(\mathfrak{U}L_0)} = \frac{\|\hat{\phi}\,\hat{u}\,N_\varnothing\|_2}{\dfrac{\sqrt{m}|\hat{\mu}_1(\omega_0)|}{\|\mu_1\|_1^{1/2}}\left\| \dfrac{\hat{\phi}\,\hat{u}\,N_\varnothing}{\hat{e}_d} \right\|_2}. \tag{37}$$

For the examples above, with $\mathcal{D}(\omega) = \dfrac{\sqrt{m}|\hat{\mu}_1(\omega_0)|}{\|\mu_1\|_1^{1/2}}\,\dfrac{1}{\hat{e}_d(\omega)}$, for $\omega\in\Omega_1$,

$\mathcal{D}(\omega_0) \cong 1$, $\mathcal{D}(\omega) \le 1$ for $|\omega| \le |\omega_0|$, and $\mathcal{D}(\omega) \ge 1$ for $|\omega| \ge |\omega_0|$. Qualitatively $\dfrac{\mathcal{SNR}(\mathfrak{U}L_s)}{\mathcal{SNR}(\mathfrak{U}L_0)} \ge 1$ for $\ell = |\omega_0|$,and $\dfrac{\mathcal{SNR}(\mathfrak{U}L_s)}{\mathcal{SNR}(\mathfrak{U}L_0)} \le 1$ for $\ell = |\omega_0| + \Delta$ and for Δ sufficiently large. This example highlights the difference between the two independent statements $\mathfrak{U}L_s \ge \mathfrak{U}L_0$ and $\mathcal{SNR}(\mathfrak{U}L_s) \ge \mathcal{SNR}(\mathfrak{U}L_0)$.

The elementary $\mathfrak{D}$ makes this example a candidate for the use of $\hat{e}_d$ as the transfer function of a linear operator which normalizes L_s to have the same noise power spectral density as L_0, as was mentioned at (29). Let $\mathfrak{E}_d$ denote this linear operator. Then, proceeding exactly as for (37)

$$\frac{\mathcal{SNR}(\mathfrak{U}\mathfrak{E}_d L_s)}{\mathcal{SNR}(\mathfrak{U}L_0)} = \frac{\hat{e}_d(\omega_0)}{\dfrac{\sqrt{m}\,|\hat{\mu}_1(\omega_0)|}{\|\mu_1\|_1^{1/2}}} \quad , \; \omega_0 \in \Omega_1 \; . \tag{38}$$

For the cases considered this is approximately unity except for ω_0 near the zero set of $\hat{\mu}_1$. The significance of (38) is that it is an explicit example of Observations 1 and 2: on Ω_1 L_s is at least as good as L_0 with the additional feature of extended frequency response outside of Ω_1. That is, on Ω_1 there is no penalty for the additional response outside of Ω_1. On the other hand, (38) only gives essentially equivalent performance on Ω_1, despite the fact that $\hat{e}_d$ has no zeroes.

Once again, any advantage due to this later depends on

$$\sup_{\mathfrak{B}} \frac{\mathcal{SNR}(\mathfrak{U}\mathfrak{B}\mathfrak{E}_d L_s)}{\mathcal{SNR}(\mathfrak{U}\mathfrak{E}_d L_s)} = \sup_{\hat{\mathfrak{b}}} \frac{\| \hat{\phi}\, \hat{\mathfrak{u}}\, N_\emptyset \|_2}{\left\| \dfrac{\hat{\phi}\, \hat{\mathfrak{u}}\, \hat{\mathfrak{b}}\, N_\emptyset}{\hat{\mathfrak{b}}(\omega_0)} \right\|_2} , \tag{39}$$

where $\mathfrak{B}$ is a linear operator and $\hat{\mathfrak{b}}$ its transfer function. It is clearly possible for $\mathfrak{E}_d L_s$ to be the optimal: for example, consider $\hat{\phi}\hat{\mathfrak{u}}N_\emptyset$ constant and $\hat{\mathfrak{b}}$ convex.

This consideration is of significance for operators $\mu_1, \mu_2, \ldots, \mu_m$ that are not characteristic functions of sets but rather have each $\hat{\mu}_i$ approximately compactly supported. The primary example here is the diffraction limited lens. If strongly coprime convolvers $\mu_1, \mu_2, \ldots, \mu_m$ were such that each $|\hat{\mu}_i|$ was small outside some set, then the envelope transfer function would exhibit the same behavior. In this case, unless $\sup_{\mathfrak{B}} \mathcal{SNR}(\mathfrak{U}\mathfrak{B}\mathfrak{E}L_s)$ is substantially greater than $\mathcal{SNR}(\mathfrak{U}\mathfrak{E}L_s)$, the performance of the strongly coprime configuration will be essentially that of its constituent convolvers.

4.4 MORE COMPARISONS: STRONGLY COPRIME VERSUS CHANGE OF SCALE

Let L_s be the same as above. In the above L_s was compared with L_0, where L_0 was chosen to be μ_1 and $N_1^2 = \|\mu_1\|_1 N_\emptyset^2$. In these cases μ_1 was the "best" in the sense $\Omega_i \subset \Omega_1$ $i=1,2,\ldots,m$. Here L_s will be compared with a one parameter family of such L. Define L_Δ by the trivial configuration of m parallel, identical $\mu_{<\Delta>}$ as in (10), where $N_\Delta^2 = \|\mu_{<\Delta>}\|_1 N_\emptyset^2$ and $\mu_{<\Delta>}(x) = \mu_1\left(\frac{x}{\Delta}\right)$, $\Delta > 0$.

The primary result of this section is

OBSERVATION FOR FIXED NUMBER OF CHANNELS: Fix the number of parallel convolvers in both L_s and L_Δ to be m . Let the convolvers be characteristic functions of cubes on $\mathbb{R}^n$ and let the additive noise be as above. Assume that $\mathfrak{U}$ is such that $\operatorname{supp}(\hat{\mathfrak{u}}) \subset \bigcup_{j=1}^{m} \left\{\omega \in \mathbb{R}^n : \omega_i = 0,\ i\neq j\right\}$. Then for $n \geq 2$

$$\mathfrak{U}L_s \geq \mathfrak{U}L_\Delta \qquad \text{for all } 0 < \Delta \leq 1 . \tag{40}$$

COROLLARY TO OBSERVATION: For the conditions in the Observation above, it is advantageous to construct L_s using sets that are as large as possible.

APPLICATION OF THE COROLLARY: In parallel scanned imaging systems with square detectors wherein the systems are ranked using some $\mathfrak{U}$

meeting the conditions of the Observation (e.g., horizontal or vertical bars), the detector size should be sufficiently large so that the array of detectors fills the image, and the detector sizes in the array should constitute a strongly coprime collection. (This application depends on sufficiently high sampling rates. See Chapter 3.)

The Observation is illustrated in Figure 21 for n = 2. For Figure 21 L_s is as in Figure 20: in the notation just above L_s is configured from the parallel convolvers $\mu_{<1>}$, $\mu_{<\sqrt{2}>}$, $\mu_{<\sqrt{3}>}$, and μ_1 is the characteristic function of the unit square. For this L_s the envelope transfer function $\hat{e}_d$ is compared with $\sqrt{3}\,\dfrac{|\hat{\mu}_{<\delta>}|}{\|\mu_{<\delta>}\|_1^{1/2}}$, as in (35), for δ = 1, 0.5, 0.2, and 0.1 . For $\delta = 1$ see Figure 20; for δ = 0.5, 0.2, and 0.1 see Figure 21. The observation in (40) is clearly evident. (Here we neglect *Observation 3* of the last section by means of a broad interpretation of $\cong$ in *Observation 1.*)

The Observation (40) depends on the following properties. The first, which is again an approximation, is that for $A_j = \left\{\omega \in \mathbb{R}^n : \omega_i = 0,\ i\neq j\right\}$, the ω_j-axis,

$$\hat{e}_d\big|_{A_j}(\omega) \cong C|\omega|^{-1} . \tag{41}$$

The second is that $\dfrac{|\widehat{\mu_{<\delta>}}(\omega)|}{\|\mu_{<\delta>}\|_1^{1/2}} = \delta^{n/2}\dfrac{|\hat{\mu}_1(\delta\omega)|}{\|\mu_1\|_1^{1/2}}$. Hence, for $n \geq 2$, for $\delta \leq 1$, and for $\omega \in A_j$, $\sqrt{m}\,\dfrac{|\hat{\mu}_1(\omega)|}{\|\mu_1\|_1^{1/2}} \leq C|\omega|^{-1} \Longrightarrow \sqrt{m}\,\dfrac{|\widehat{\mu_{<\delta>}}(\omega)|}{\|\mu_{<\delta>}\|_1^{1/2}} \leq C|\omega|^{-1}$.

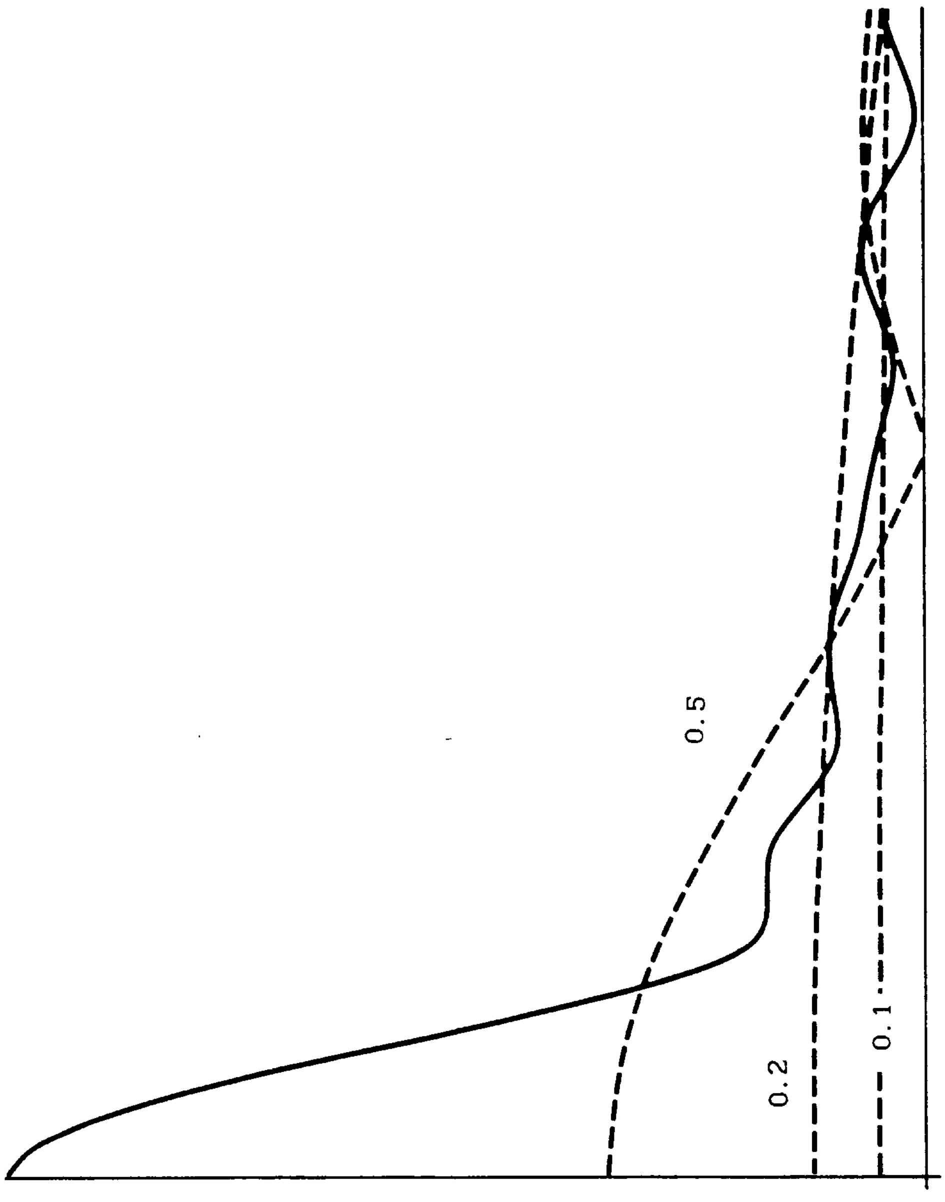


Fig. 21. $\mathbb{R}^2$; abscissa = ω_1-axis solid: $\hat{e}_d$; broken: $\sqrt{3}\,\dfrac{|\hat{\mu}_{<\delta>}|}{\|\mu_{<\delta>}\|_1^{1/2}}$, δ = 0.5, 0.2, 0.1

Figures 22 and 23 show two counterexamples for cases not addressed in the Observation. Figure 22 is for the case of the diagonal in $\mathbb{R}^2$, and Figure 23 is for $n = 1$. The Observation fails on the diagonal $\mathcal{D} = \left\{\omega=(\omega_1,\omega_2) \in \mathbb{R}^2 : \omega_1=\omega_2\right\}$ because

$$\hat{e}_d|_{\mathcal{D}}(\omega) \cong C|\omega|^{-2} . \tag{42}$$

It fails for $\mathbb{R}$ because (41) holds.

If in place of characteristic functions of cubes one uses characteristic functions of disks on $\mathbb{R}^2$, then the relationship between $\hat{e}_d$ and $\sqrt{m}\,\dfrac{|\widehat{\mu_{<\Delta>}}|}{\|\mu_{<\Delta>}\|_1^{1/2}}$ is intermediate between that of the ω_j-axis and that of the diagonal for

$$\hat{e}_d(\omega) \cong C|\omega|^{-3/2}. \tag{43}$$

The significance of the Observation (40) is that it provides a qualitative lower bound for the performance of the strongly coprime configuration. To the extent performance is characterized for the $\mathbb{L}_\Delta$, the "envelope" consisting of the collection over all Δ is a lower bound for the performance of $\mathbb{L}_s$.

All of the above has focused on performance away from the origin. If the figures are rescaled so that the $\mu_{<\Delta>}$ appear fixed with a sequence of L_s constructed from convolvers of increasing support, the Observation indicates that nothing is sacrificed away from zero while the envelope transfer function near zero is substantially increased. That is, $\mathbb{L}_s \geq \mathbb{L}_\Delta$ represents a substantial enhancement near $\omega = 0$, not merely approximately identical performance. On the other hand, this

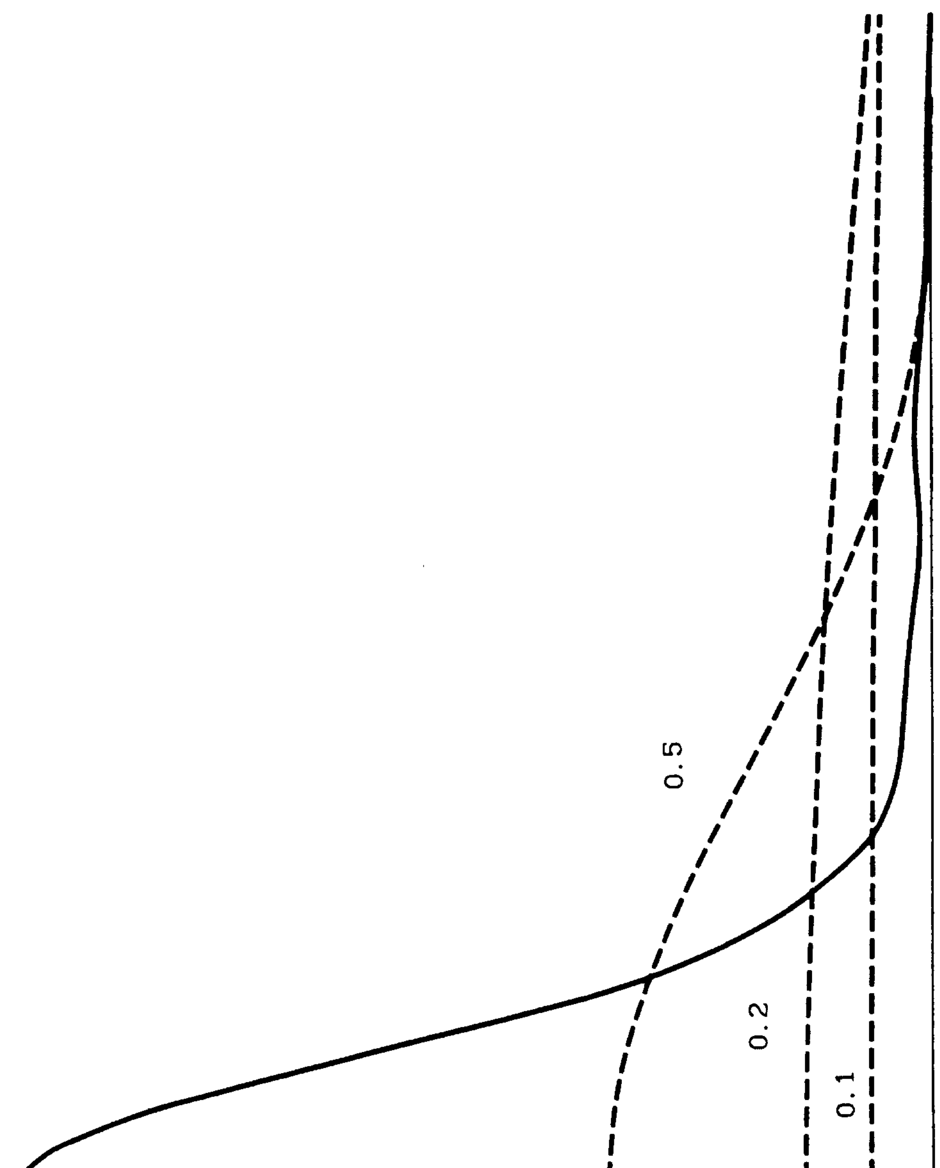


Fig. 22. R^2 ; abscissa = diagonal solid: $\hat{e}_d$; broken: $\sqrt{3}\,\dfrac{|\hat{\mu}_{<\delta>}|}{\|\mu_{<\delta>}\|_1^{1/2}}$, δ = 0.5, 0.2, 0.1

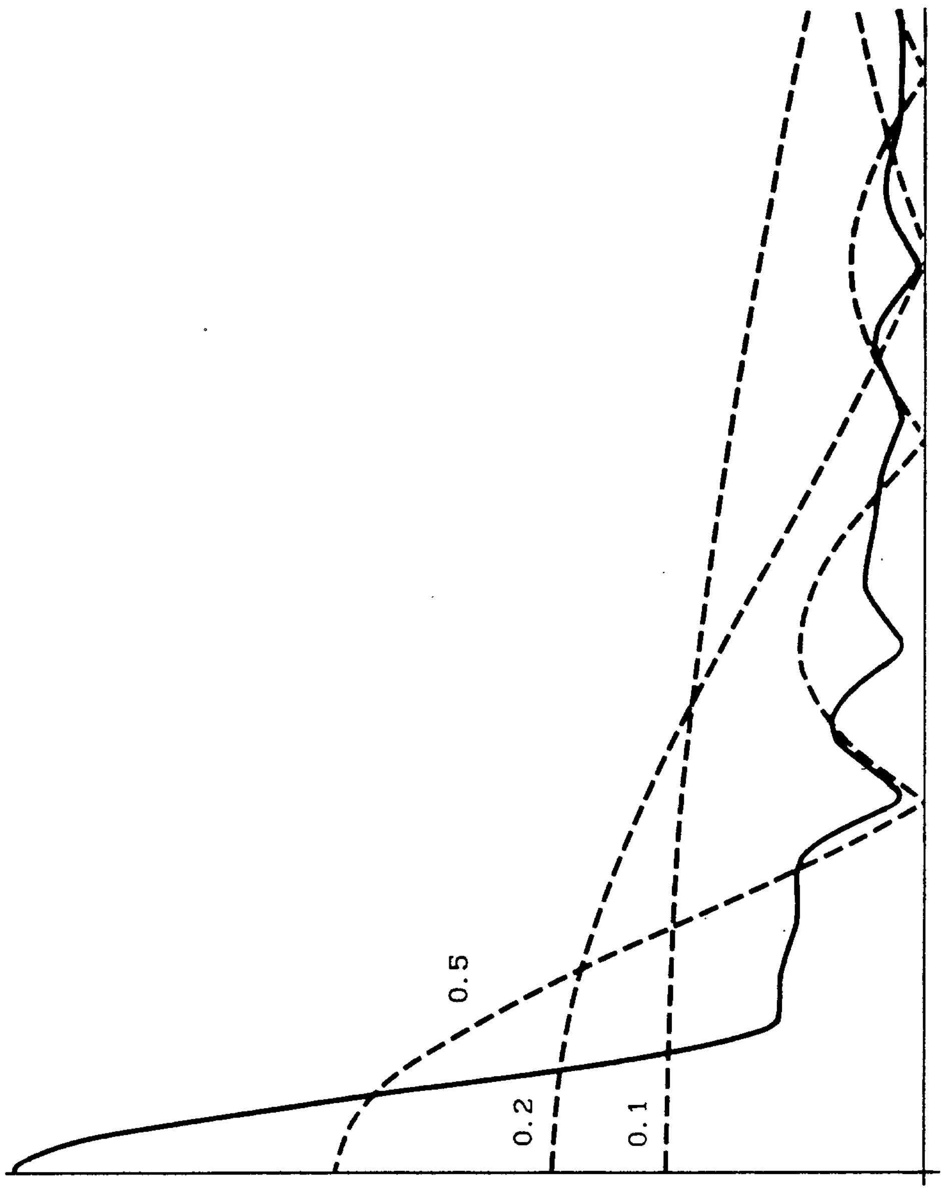


Fig. 23. $\mathbb{R}^1$; solid: $\hat{\varepsilon}_d$; broken: $\sqrt{2}\,\dfrac{|\hat{\mu}_{<\delta>}|}{\|\mu_{<\delta>}\|_1^{1/2}}$, $\delta = 0.5,\ 0.2,\ 0.1$

uniform improvement is for the case of $\mathfrak{u}$ supported by the axes. For the cases off the axes for cubes and for the case of disks there is a trade-off between some loss away from zero and the gain near zero.

BIBLIOGRAPHY

Arnold, V. I. 1973. *Ordinary Differential Equations.* Cambridge: MIT Press.

Arnold, V. I. 1978. *Mathematical Methods of Classical Mechanics.* New York: Springer-Verlag.

Berenstein, Carlos A. 1983. An application of the Andersson-Berndtsson integral representation formula. *Revue de l'Institut Elie Cartan* 8, 113-60.

Berenstein, Carlos A., P. S. Krishnaprasad, and B. A. Taylor. 1984. Deconvolution methods for multi-sensors. ARO contract DAAG29-81-D-0100, DTIC no. AD A152 351, 68 pages.

Berenstein, Carlos A., and B. A. Taylor. 1979. A new look at interpolation theory for entire functions of one variable. *Advances in Mathematics* 33, 109-43.

———. 1980a. Interpolation problems in $\mathbb{C}^n$ with applications to harmonic analysis. *J. Analyse Math.* 38, 188-254.

———. 1980b. Mean periodic functions. *Intern. J. Math. and Math. Sciences* 3, 199-236.

Berenstein, Carlos A., B. A. Taylor, and Alain Yger. 1983a. On some explicit deconvolution formulas. *Technical Digest, Signal Recovery,* Optical Society of America, Winter Meeting 1983, pp WA4-1 to WA4-4.

———— . 1983b. Sur quelques formules explicites de deconvolution. *Journal of Optics* (Paris) 14, 75-82.

Berenstein, Carlos A., and Alain Yger. 1983. Le probleme de la deconvolution. *J. Funct. Anal.* 54, 113-60.

Bishop, Richard L., and Samuel I. Goldberg. 1968. *Tensor Analysis on Manifolds.* New York: Dover.

Blicher, A. Peter. 1985. Edge detection and geometric methods in computer vision. Ph.D. diss., Rpt. No. STAN-CS-85-1041, Department of Computer Science, Stanford University, Stanford.

Collett, Thomas S., and Lindesay I. K. Harkness. 1982. Depth vision in animals. In *Analysis of Visual Behavior*, ed. D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, 111-176. Cambridge: MIT Press.

Davenport, Wilbur B., and Willian L. Root. 1958. *An Introduction to the Theory of Random Signals and Noise.* New York: McGraw-Hill.

Flanders, Harley. 1963. *Differential Forms with Applications to the Physical Sciences.* New York: Academic Press.

Federer, Herbert. 1969. *Geometric Measure Theory.* New York: Springer-Verlag.

Hörmander, Lars. 1967. Generators for some rings of analytic functions. *Bulletin Amer. Math. Soc.* 73, 943-9.

Kelleher, James J., and B. A. Taylor. 1971. Finitely generated ideals in rings of analytic functions. *Math. Annalen* 193, 225-37.

Marr, David. 1982. *Vision.* San Francisco: W. H. Freeman.

Meyer-Arendt, Jurgen R., 1984. *Introduction to Classical and Modern Optics.* 2d ed. Englewood Cliffs, N.J.: Prentice-Hall.

O'Neill, Barrett. 1983. *Semi-Riemannian Geometry with Applications to Relativity.* New York: Academic Press.

Prazdny, K. 1983. On the information in optical flows. *Computer Vision, Graphics, and Image Processing* 22, 239-59.

Ratches, James A., Walter R. Lawson, et al. 1975. Night Vision Laboratory Static Performance Model for Thermal Viewing Systems, U.S. Army Electronics Command Technical Report, ECOM-7043.

Sternberg, Shlomo. 1983. *Lectures on Differential Geometry.* 2d ed. New York: Chelsea.

Warner, Frank W. 1971. *Foundations of Differentiable Manifolds and Lie Groups.* Glenview, Ill.: Scott, Foresman.

Wheeden, Richard L., and Antoni Zygmund. 1977. *Measure and Integral, An Introduction to Real Analysis.* New York: Marcel Dekker.

CURRICULUM VITAE

Name: Emil Vincent Patrick

Permanent Address: [redacted]

Degree and date to be conferred: Ph.D., 1987

Date of birth: [redacted]

Place of birth: Sewickley, PA

Secondary education: Ambridge High School, Ambridge, PA 1963

Collegiate institutions attended:	Dates	Degree	Date of Degree
Thiel College	1963-1967	B.A.	1967
University of Maryland	1978-1982	M.A.	1982
University of Maryland	1982-1987	Ph.D.	1987

Major: B.A. Physics; M.A., Ph.D. Mathematics

Minor: Philosophy

Professional publications:

1. "Identification of the sources of signal induced noise in microchannel plates," Night Vision Laboratory Report, 1970.
2. "Microchannel plates with high open area ratio," (with A. Asam), *Proc. Electro-Optical Systems Design Conference-1976/International Laser Expo.*, September 1976, 629-635.
3. "Microchannel plate in wall fabrication, method and apparatus," (with L.C. Spessard, K. Villhauer), U.S. Patent 4,198,225; April 15, 1980.
4. "Vacuum interlock," (with H.K. Dickson, H.L. Dunmire), U.S. Patent 4,212,317; July 15, 1980.
5. "Vacuum sealed manipulator," (with H.K. Dickson), U.S. Patent 4,212,575; July 15, 1980.
6. "Method and apparatus to fabricate image intensifier tubes," (with H.K. Dickson, H.L. Dunmire), U.S. Patent 4,286,833; September 1, 1981.
7. "Selection of dimensions in image intensifiers," Night Vision Laboratory Report, 1984.
8. "Nonlinear gain microchannel plate image intensifier tube," (with E.J. Bender, R.E. Franseen), U.S. Patent 4,625,106; November 25, 1986.
9. "Wide angle and graded acuity intensifier tubes," U.S. Patent application SN 935,359; November 26, 1986.

Professional positions held:

1986-1987	Graduate Research Assistant, SRC Fellowship System Research Center University of Maryland College Park, MD 20742
1968-1986	Physicist Center for Night Vision & Electro-Optics Ft. Belvoir, VA 22060
1978-1978	Teaching Assistant Mathematics Department University of Maryland College Park, MD 20742